22 MAR 1963

# PERFORMANCE CRITERIA FOR LINEAR
# CONSTANT-COEFFICIENT SYSTEMS WITH RANDOM INPUTS

TECHNICAL DOCUMENTARY REPORT ASD-TDR-62-470

January 1963

Flight Control Laboratory
Aeronautical Systems Division
Air Force Systems Command
Wright-Patterson Air Force Base, Ohio

Project No. 8219, Task No. 821904

## NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Qualified requesters may obtain copies of this report from the Armed Services Technical Information Agency, (ASTIA), Arlington Hall Station, Arlington 12, Virginia.

This report has been released to the Office of Technical Services, U.S. Department of Commerce, Washington 25, D.C., in stock quantities for sale to the general public.

Copies of this report should not be returned to the Aeronautical Systems Division unless return is required by security considerations, contractual obligations, or notice on a specific document.

B

Aeronautical Systems Division, Dir/Aeromechanics,
Flight Control Lab, Wright-Patterson AFB, Ohio.
Rpt Nr ASD-TDR-62-470, PERFORMANCE CRITERIA FOR
LINEAR CONSTANT-COEFFICIENT SYSTEMS WITH RANDOM
INPUTS. Final report, Jan 63. 128p. incl illus.,
tables, 56 refs.
                              Unclassified Report

Performance measures and criteria for linear
constant-coefficient systems with random inputs are
critically surveyed, with particular reference to
"pencil and paper" methods of optimizing flight
control systems. For stationary inputs many
criteria are shown to be equivalent to minimum mean
square error. This criterion is easy to apply, but
yields lightly damped systems and is unselective,
i.e., the mean square error of a wide variety of
off-optimum systems is little higher than that of
the optimum. It is concluded that no criterion yet

( over )

1. Flight control
2. Linear constant-
   coefficient systems
3. Flight control systems
4. Flight control
   optimization
5. Stochastic processes
I. AFSC Project 8219,
   Task 8219004
II. Contract No.
    AF 33(616)-7841
III. Systems Technology,
     Inc., Inglewood, Calif.
IV. Raymond Magdaleno
    Julian Wolkovitch
V. STI TR-121-1

proposed is suitable as a sole criterion for flight
control system optimization.

VI. Aval fr OTS
VII. In ASTIA collection

Aeronautical Systems Division, Dir/Aeromechanics,
Flight Control Lab, Wright-Patterson AFB, Ohio.
Rpt Nr ASD-TDR-62-470, PERFORMANCE CRITERIA FOR
LINEAR CONSTANT-COEFFICIENT SYSTEMS WITH RANDOM
INPUTS. Final report, Jan 63. 128p. incl illus.,
tables, 56 refs.
                              Unclassified Report

Performance measures and criteria for linear
constant-coefficient systems with random inputs are
critically surveyed, with particular reference to
"pencil and paper" methods of optimizing flight
control systems. For stationary inputs many
criteria are shown to be equivalent to minimum mean
square error. This criterion is easy to apply, but
yields lightly damped systems and is unselective,
i.e., the mean square error of a wide variety of
off-optimum systems is little higher than that of
the optimum. It is concluded that no criterion yet
( over )

1. Flight control
2. Linear constant-
   coefficient systems
3. Flight control systems
4. Flight control
   optimization
5. Stochastic processes
I. AFSC Project 8219,
   Task 8219004
II. Contract No.
    AF 33(616)-7841
III. Systems Technology,
     Inc., Inglewood, Calif.
IV. Raymond Magdaleno
    Julian Wolkovitch
V. STI TR-121-1

proposed is suitable as a sole criterion for flight
control system optimization.

VI. Aval fr OTS
VII. In ASTIA collection

## FOREWORD

This report represents one phase of an effort directed at the use of performance criteria as elements in flight control system optimization studies. The research reported was sponsored by the Flight Control Laboratory of the Aeronautical Systems Division under Project No. 8219. It was conducted at Systems Technology, Inc., under Contract No. AF 33(616)-7841. A companion report (Ref. 1) dealing with performance criteria for deterministic inputs has already been published under this contract. The ASD project engineers were Mr. R. O. Anderson and Lt. L. Schwartz of the Flight Control Laboratory. The principal contributors to the report are listed as authors.

The authors wish to express their thanks to Messrs. D. T. McRuer and Dunstan Graham, principal investigators, who planned the general approach followed in both reports and contributed many details. Thanks are also due to Lt. L. Schwartz for his thorough check of and valuable comments on the report, and to Mr. A. V. Phatak for contributing many of the calculations. Acknowledgment is gratefully made to Messrs. J. Taira and R. N. Nye and Misses N. Crawford and D. Lewis for their careful work in preparing the report.

ABSTRACT

A critical survey and assessment has been made of performance measures
and associated criteria for linear constant-coefficient systems with random
inputs. The suitability of each criterion for flight control system
optimization using "pencil and paper" methods has been investigated by con-
sidering its validity, selectivity, and ease of application. Simplifications
are introduced by substituting for actual flight control systems lower order
equivalent systems having similar dynamics, and by replacing certain random
inputs and criteria with transient analogs and "compatible" deterministic
criteria. The latter simplification enables the calculation of random input
performance measures to be replaced by calculations involving more easily
visualized deterministic quantities.

It is shown that, for stationary inputs, a wide variety of criteria
reduce to minimum mean square error. This criterion is easy to use and
has a compatible deterministic form, but yields lightly damped systems and
is unselective (i.e., the mean square error of a wide variety of off-optimum
systems is little higher than that of the optimum). However, no criterion
was found that had the advantages of minimum mean square error without
comparable or worse disadvantages. For certain nonstationary problems,
the probabilistic square error criterion appears promising. Numerous
improved techniques for evaluating criteria and several examples are
presented. It is concluded that none of the criteria that have been pro-
posed to date are suitable as sole criteria for flight control system
optimization.

PUBLICATION REVIEW

This report has been reviewed and is approved.

FOR THE COMMANDER

*C. B. Westbrook*

C. B. WESTBROOK
Chief, Aerospace Mechanics Branch
Flight Control Laboratory

Page

ILLUSTRATIONS

Page

# NOMENCLATURE

| | |
|---|---|
| a | A parameter used to obtain the Wiener system |
| a | Input power spectrum break frequency |
| $a_k$ | A random variable with a known Gaussian joint probability density function |
| $a_n$ | Loss function coefficient |
| a(t) | A general stationary random signal |
| A | See Eq 207 |
| $A_0$, $A_1$ | Denominator coefficients of a general second-order system transfer function (see Eq 96) |
| A(s) | See Eq 88 |
| $A = A(j\omega)$ | See Eq 201 |
| b(t) | A general stationary random signal |
| $B_0$, $B_1$ | Numerator coefficients of a general second-order system transfer function (see Eq 96) |
| B(s) | See Eq (88) |
| $B = B(j\omega)$ | See Eq 201 |
| c | System output c(t) |
| $c_d$ | Desired system output $c_d(t)$ |
| $c_g$ | Output due to system having weighting function g(t) |
| $c_n$ | A constant |
| $c_t$ | Transient analog (Chapter IV) of c |
| C(s) | $= \mathcal{L}\left[c(t)\right]$ |
| $c_W$ | Output of the Wiener system |
| e | $= 2.712 \cdots$ |
| erfc (x) | The complementary error function $= 1 - \text{erf}(x)$ |
| E | Expected value |
| E(s) | $= \mathcal{L}\left[\epsilon(t)\right.$ |
| $E_i(s)$ | $= \mathcal{L}\left[\epsilon_i(t)\right]$ |

| | |
|---|---|
| $E_t(\tau)$ | Generalized error function for deterministic inputs |
| $EMS_W$ | Frequency-weighted mean square error |
| $f$ | Frequency, $= \omega/2\pi$ |
| $f$, $f(t)$ | Signal or command portion of the input |
| $f(r)$ | Nonlinear gain function operating upon input |
| $f(\epsilon)$ | Benedict and Sondhi's loss function |
| $f_d(t)$ | See Glover's criterion and Table I |
| $f_k(t)$ | A known deterministic function of time |
| $\mathcal{F}$ | Fourier transform operator |
| $g(t)$ | (i)  A general weighting function |
| | (ii)  A deterministic signal representable as a polynomial in $\tau$ |
| $G$ | Open-loop transfer function |
| G.E.F. | Generalized error function |
| $h$, $h(t)$ | General transfer function |
| $h(x)$ | See Eq 114 |
| $h_W(t)$ | Weighting function of the Wiener system |
| $H = H(s)$ | A transfer function |
| $H_a(s)$, $H_b(s)$ | See Eq 167 |
| $H_1(s)$ | A transfer function |
| $H_{EMS_W}(s)$ | Optimum frequency-weighted mean square error system |
| $H_p(s)$ | Transfer function of the Phillips system |
| $H_W(s)$ | Transfer function of the Wiener system |
| $H_X(s)$ | See Eq 204 |
| $IE^2$ | Integrated error squared, $\int_o^\infty \epsilon^2 dt$ |
| ITAE | Integrated time moment of absolute error, $\int_o^\infty t\lvert\epsilon\rvert dt$ |
| $ITE^2$ | Integrated time moment of error squared, $\int_o^\infty t\epsilon^2 dt$ |

| | |
|---|---|
| $j$ | $= \sqrt{-1}$ |
| $J(s)$ | See Eq 89 |
| $k(t)$ | Optimum transfer function defined by Eq 106 |
| $k_i(t)$ | An "ideal" transfer function defined by Eq 67 |
| $k_n$ | Constant defining optimum weighting function for minimum probabilistic square error |
| $k_{1,2}$ | See Eq 120 |
| $K$ | Input amplitude descriptor (see Fig. 17) |
| $K_m K_e$ | Product of controller gains |
| $L$ | (i)  A specified level of output or effort<br>(ii)  A loss function, $L(\epsilon)$ |
| $L(s)$ | See Eq 89 |
| $\mathcal{L}$ | Laplace transform operator |
| $m(t)$ | A stationary random signal with zero mean |
| $M_D$ | A deterministic input performance measure |
| $M_R$ | A random input performance measure |
| MWFE | Mean weighted function of error, see Eq 74 |
| MWSE | Mean weighted square error, see Eq 73 |
| $n$ | An exponent |
| $n, n(t)$ | Noise or unwanted component of the input |
| $n_t(t)$ | Transient analog of random noise |
| $N$ | (i)  Largest exponent in Kaufman's performance measure<br>(ii)  Noise amplitude |
| $N_M$ | Average number of maxima per second of a stationary random quantity |
| $N_0$ | Average number of zero crossings per second of a stationary random quantity |
| $N_{\epsilon 0}$ | Average number of zero crossings per second of effort |

| | |
|---|---|
| p | Probability criterion (page 55) |
| $p(a)$ | Probability density function of a |
| $p(a,b)$ | Joint probability density function of two quantities, a and b |
| $p(a/b)$ | $= \dfrac{p(a,b)}{p(b)}$, conditional probability density function of a, given b |
| $p_i$ | Zeros of $L(s)$ |
| $p_2$ | Second probability density function |
| $p(t)$ | Weighting function for end-sigma performance measure |
| P.E. | Glover's performance measure, $\displaystyle \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} \frac{\epsilon^2(t)}{[c_d(t)]^2 + \delta^2} \, dt$ |
| r | An input |
| $\hat{r}$ | A sequence of past input values, $r_0, r_1, \cdots r_n$ |
| $r_i$ | A sampled value of the input |
| $r_t(t)$ | Transient analog of a random signal $r(t)$ |
| $r_D$ | A deterministic input |
| $r_R$ | A stationary random input |
| $R_{ab}(\tau)$ | $= \displaystyle \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} a(t)b(t + \tau)dt$, crosscorrelation between two stationary random signals, $a(t)$ and $b(t)$ |
| $\left.\begin{array}{l} R_m \\ R_n \end{array}\right\}$ | Special autocorrelation functions defined in Eq 71 |
| $R(s)$ | Laplace transform of input |
| $R_t(s)$ | $= \mathcal{L}[r_t(t)]$ |
| $R_{Wab}$ | Weighted crosscorrelation or autocorrelation (a = b) function, see Eq 116 |
| s | Laplace transform complex variable, $\sigma + j\omega$ |
| s, s(t) | Signal or command portion of the input |
| $s_t(t)$ | Transient analog of random signal |

$S = S(s)$ $\qquad = \mathcal{L}\left[s_t(t)\right]$

$\mathcal{S}_{X_1 X_1}(\omega)$ $\qquad = X_1(j\omega)X_1(-j\omega)$ energy spectral density

$t$ $\qquad$ Time

$t_d$ $\qquad$ Delay time

$t_i$ $\qquad$ Time of application of a step input

$T$ $\qquad$ (i)     A fixed time

$\qquad\qquad\qquad$ (ii)    Sampling period

$\qquad\qquad\qquad$ (iii)   Observation time

$u$ $\qquad$ A dummy time variable

$u(t)$ $\qquad$ A step function commencing at $t = 0$

$v$ $\qquad$ General signal break frequency

$w(t)$ $\qquad$ A time-weighting function

$W$ $\qquad = W(\omega)$ (see below)

$W(t)$ $\qquad$ A time-weighting function

$W(\omega)$ $\qquad$ Frequency weighting function

$W_a(\omega)$ $\qquad$ See Eq 224

$W_b(\omega)$ $\qquad$ See Eq 225

$x$ $\qquad$ A dummy time variable

$x_1(t), x_2(t)$ $\qquad$ Functions of time

$X_1(s)$ $\qquad = \mathcal{F}\left[x_1(t)\right]$

$X_2(s)$ $\qquad = \mathcal{F}\left[x_2(t)\right]$

$y$ $\qquad$ A dummy time variable

$Z$ $\qquad$ See Eq 59

$\alpha$ $\qquad$ See Eq 211

$\alpha_i$ $\qquad$ A system adjustable parameter

$\alpha_k$ $\qquad$ See Eq 39

$\alpha_0, \alpha_1, \alpha_2$ $\qquad$ Weighting parameters

| | |
|---|---|
| $\beta$ | See Eq 211 |
| $\gamma_r$ | A Lagrange multiplier |
| $\Gamma(x)$ | Gamma function of x |
| $\delta$ | (i) A constant in Glover's criterion indicating the lowest accuracy of interest |
| | (ii) Unit impulse function |
| $\epsilon$ | (i) System error (difference between actual and desired output) |
| | (ii) Actuating error (difference between input and output) |
| $\epsilon_a(t)$, $\epsilon_b(t)$ | Transient signals used to generate the transient analog of the cross power spectra in Chapter IV |
| $\epsilon_n$ | Random component of error |
| $\epsilon_W$ | Frequency-weighted error signal |
| $\epsilon_1(t)$ | Transient analog of random signal |
| $\epsilon_2(t)$ | Transient analog of random noise |
| $\epsilon_{s0}$ | Systematic error at time $t_0$, i.e., that part of the error response due to the deterministic component of the input |
| $\zeta$ | Damping ratio |
| $\eta$ | Prediction or interpolation time |
| $\eta_1$ | See Eq 42 |
| $\theta_a(r)$ | Orthogonal polynomial in r of degree a used in synthesizing Lubbock's optimum system |
| $\kappa$ | Constant |
| $\mu_\nu$ | $\nu^{\text{th}}$ moment of impulsive response as defined in Eq 65 |
| $\nu$ | A system parameter |
| $\rho_{aa}$ | Normalized autocorrelation function, $\rho_{aa}(\tau) = \dfrac{R_{aa}(\tau)}{\sigma_a^2}$ |
| $\sigma^2$ | Mean square value of some quantity |
| $\sigma_{ab}$ | Normalized crosscorrelation function, $\sigma_{ab}(\tau) = \dfrac{R_{ab}(\tau)}{\sigma_a \sigma_b}$ |

$\varsigma$  (i)  End-sigma performance measure

$$\int_{0}^{\infty} E\big[F \ \epsilon(t), \ t, \ \nu, \ \nu_1, \ \nu_2, \ \cdots \nu_n \ p(t)\big] dt$$

(ii)  Probabilistic square error

$$\int_{0}^{\infty} \big[\epsilon(t)\big]^2 p(t) dt$$

| | |
|---|---|
| $\tau$ | A dummy time variable |
| $\tau, \ \tau_1, \ \tau_2$ | Time delays |
| $\Phi(\omega)$ | Power spectral density |
| $\Phi_{rr}(\omega)$ | Input power spectrum, $= \Phi_{rr}^{+}(j\omega)\Phi_{rr}^{-}(j\omega)$ |
| $\Phi_{\epsilon\epsilon}(\omega)$ | Error power spectrum |
| $\Phi_{ab}(j\omega)$ | A general cross-spectrum |
| $\chi$ | Kaufman's performance measure, $\displaystyle\sum_{n=1}^{N} c_n \overline{\epsilon^{2n}}$ |
| $\psi$ | A function of present and past input samples |
| $\psi_{ii}, \ \psi_{ir}, \ \psi_{rr}$ | See Eq 101 |
| $\omega$ | Frequency, imaginary part of Laplace transform complex variable, $s = \sigma + j\omega$ |
| $\overline{\phantom{xxx}}$ | Time average |
| $\sim\!\sim\!\sim$ | Ensemble average |
| $\mid \ \mid$ | Absolute value |
| $[\ ]^*$ | Complex conjugate of $[\ ]$ |
| $[\ ]^+$ | The ratio of the left-half-plane poles and zeros of the expression in brackets (see Eq 19) |
| $[\ ]^-$ | The ratio of the right-half-plane poles and zeros of the expression in brackets (see Eq 19) |
| $[\ ]_+$ | Expand the expression in brackets in partial fractions and then keep only fractions with left-half-plane poles (see Eq 19) |
| $\doteq$ | Approximately equals |
| $\equiv$ | Identically equals |

# INTRODUCTION

## A. ARRANGEMENT AND CONTENTS OF THE REPORT

The reader who wishes to obtain a more complete knowledge of random input performance criteria than can be gleaned from the Abstract and Conclusions alone, but who lacks the time or inclination to read the whole report, is advised to read the remainder of this Introduction, the summaries on pp. 27, 28, and 63, and all of the concluding chapter.

This report presents the results obtained during the second phase of a generalized study of dynamic performance measures for automatic flight control systems. Such systems are subject to both random and deterministic inputs, and it has been found convenient to present the results of the study in two parts, Ref. 1 dealing with deterministic inputs and the present report with random inputs. This introductory section defines some important terms used throughout the report, and sets out the viewpoint from which the various performance criteria will be assessed. The report consists principally of a survey and critical assessment of published performance criteria. The number and diversity of the criteria examined is so great that it would take many pages of text to summarize the results of the study in such a fashion that something is said about each criterion. Table I has, therefore, been prepared: it lists the criteria, states the applicable input conditions, and briefly summarizes the assessment of each criterion given in the main text of the report. In order to keep the size of Table I within reasonable bounds, the criteria definitions, etc., are terse; fuller explanations will be found on the pages indicated in the table.

Many criteria prove to be equivalent or closely related to other criteria. These relationships are compactly illustrated by "family trees" given on pp. 27, 28, and 63.

Some familiarity is presumed on the part of the reader with the elements of random process control theory. Definitions of such standard terms as "stationary" and "Gaussian" are given in Ref. 2 and elsewhere, and it is not thought necessary to repeat them here. It is also desirable (though not essential) that the reader have some acquaintance with Wiener and Phillips-type optimization because only a brief review of these well-known topics is given. Unless otherwise

---

stated, it may be assumed that random inputs referred to are stationary and that the systems are time-invariant and that the ergodic hypothesis is applicable.

B.  DEFINITIONS OF "PERFORMANCE MEASURE," "PERFORMANCE CRITERION," "LOSS FUNCTION," AND "EQUIVALENT SYSTEM"

Performance Measure:  A quantity characterizing some aspect of dynamic performance, such as stability, response to desired inputs, response to unwanted inputs, accuracy, etc.

Performance Criterion:  A standard or reference value of some performance measure.  It provides a basis for a rule or test by which some aspect of dynamic performance is evaluated in forming a judgment of system quality.

Loss Function:  A function only of the controlled quantity, $c(t)$, or of the instantaneous error, $\epsilon(t)$, between the desired and actual values of $c(t)$.  Its magnitude at any given $\epsilon(t)$ $|$or $c(t)|$ indicates the importance attached to that $\epsilon(t)$ $|$or $c(t)|$.  For example, in Fig. 1 the loss function is appropriate to a
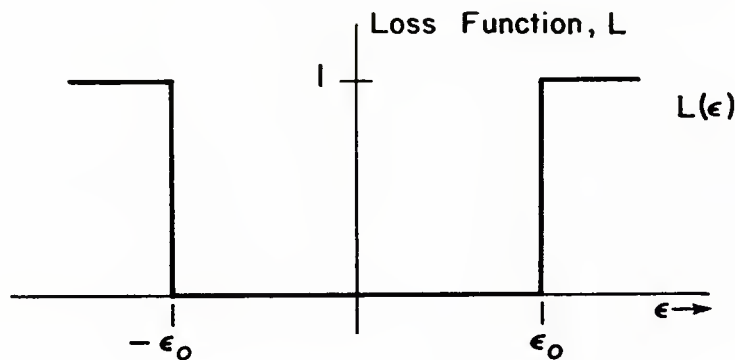


Figure 1.  Typical Loss Function

situation where all errors greater than $|\epsilon_0|$ are of equal importance, while errors of smaller magnitude are of zero importance.  Most of the performance measures discussed in this report are time-averaged loss functions of error, $\overline{L(\epsilon)}$ (for example, $\overline{\epsilon^2}$).

Equivalent System:  A simpler system which has closed-loop dynamic characteristics approximately equivalent to those of some actual system.  It normally has a transfer function of first to fourth order plus, if necessary, time delay terms to represent high-frequency leads and lags.

The general procedure by which equivalent systems are derived is most clearly illustrated by an example taken from Ref. 1 and repeated here for ease of reference.

2

TABLE I

SUMMARY OF RANDOM INPUT PERFORMANCE MEASURES
For Linear Constant-Coefficient Systems

| MEASURE | SEE PAGE | TYPE OF SYSTEM INPUTS | ASSOCIATED CRITERION | CRITERION ASSESSMENT | REMARKS |
|---|---|---|---|---|---|
| Mean square error $$\overline{\epsilon^2} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \epsilon^2(t)dt$$ | 17 | Stationary random desired and unwanted signals (Ref. 3 and 4). | Minimum value. System form may be fixed (Phillips-type optimization) or free (Wiener optimization). For Wiener optimization, system form is implicitly determined by form of input and desired response spectra. | $\overline{\epsilon^2}$ tends to be unselective and the optimum system not well damped. If input signals are stationary and Gaussian, the Wiener optimum linear system is the absolute optimum of all filters, linear or nonlinear. | Exceptionally well developed theory and application (Ref. 3, 5, 6, 7, and 8). Primary difficulty in flight control application is to find adequate expressions for desired and unwanted signals. $\overline{\epsilon^2}$ can be expressed analytically in terms of system and signal parameters; hence, the effect of parameter variations on $\overline{\epsilon^2}$ can be assessed in a straightforward manner. In practice, high-order systems yield complicated expressions, and trial and error is used to supplement the analytical treatment (Ref. 6). |
| | 119 | Stationary random desired signal plus constraints on system. | Example: Minimum $\overline{\epsilon^2}$ while $$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} (torque)^2 dt = const.$$ (Ref. 9 and 10) or minimum weighted plant input power with $\overline{\epsilon^2} \leq$ constant (Ref. 10 and 11) or other system constraints. | No general assessment due to wide variety of possible constraints. | Well-developed theory and application using Lagrange multiplier techniques (Ref. 6, 9, 10, and 11). |
| | 57 | Command signal composed of a polynomial time function of order n, plus a stationary random signal plus an unwanted input which is a stationary random signal. | Minimum $\overline{\epsilon^2}$ [first n moments of $\mathcal{L}^{-1} E(s)/R(s)$ specified] (Ref. 7 and 12). | Even simple problems require considerable computational effort. | Well-developed theory combining error coefficients with random input system theory (Ref. 7 and 12). |
| Mean function of error $$\overline{f(\epsilon)} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} f(\epsilon)dt$$ | 26 | Stationary Gaussian random desired and unwanted signals. | Minimum value. System form (i) free or (ii) fixed. | Yields minimum $\overline{\epsilon^2}$ system [(i) Wiener or (ii) Phillips]. | $f(\epsilon) = |\epsilon|^{n_1} + |\epsilon|^{n_2} + \cdots |\epsilon|^{n_k}$. Each n must be positive, but need not be an integer. |
| | 31 | Stationary random desired and unwanted signals. | Minimum value. System form free. | A) For stationary Gaussian processes, yields the Wiener system. B) If signals are nonstationary Gaussian, the optimum system is a time-varying linear system which also minimizes $\overline{\epsilon^2}$. C) For non-Gaussian (stationary or nonstationary) signals, the general solution is not known. | Sherman's (Ref. 13) loss function, $f(\epsilon)$, is defined as $f(0) = 0$     null property; $f(\epsilon) = f(-\epsilon)$     symmetry property; $f(\epsilon_1) \geq f(\epsilon_2) \geq 0$    monotonic property for $\epsilon_1 > \epsilon_2 > 0$. The determination of the optimum system also depends on some properties of the input distribution (see text). |
| | 33 | Stationary Gaussian random desired and unwanted signals. | Minimum value. System form (i) free or (ii) fixed. | Yields minimum $\overline{\epsilon^2}$ system [(i) Wiener or (ii) Phillips]. | Chang's (Ref. 10) loss function is the same as Sherman's (Ref. 13) above, although it may be asymmetric if in addition $\overline{\epsilon} \equiv 0$. Brown (Ref. 15) also considered an asymmetric loss function. |
| | 34 | Gaussian random desired (Ref. 14) and unwanted signals plus additional signal terms of the form $\sum_k a_k f_k(t)$. $f_k(t)$'s are known deterministic functions $a_k$'s are random variables with a known Gaussian joint probability density function | Minimum value. System form free. | The absolute optimum system is the minimum $\overline{\epsilon^2}$ linear (possibly time-varying) system plus a bias term on the output. | The loss function is an arbitrary function of the error. If it is the same as Sherman's (Ref. 13) loss function (see above), then the output bias term is zero. |

TABLE I (Continued)

SUMMARY OF RANDOM INPUT PERFORMANCE MEASURES
For Linear Constant-Coefficient Systems

| MEASURE | SEE PAGE | TYPE OF SYSTEM INPUTS | ASSOCIATED CRITERION | CRITERION ASSESSMENT | REMARKS |
|---|---|---|---|---|---|
| Kaufman's performance measure (Ref. 16) $$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \sum_{n=1}^{N} c_n \epsilon^{2n} dt$$ | 33 | Stationary random desired and unwanted signals. | Minimum value (Ref. 16). System order specified. | Yields the same system as $\overline{\epsilon^2}$ for Gaussian inputs. Very few results for non-Gaussian inputs. | Original reason for this criterion was to handle non-Gaussian signals. Previous work (Ref. 16) considered only linear systems with inputs having probability density functions symmetric about the mean. In some cases, Kaufman's performance measure is minimized by minimizing $\overline{\epsilon^2}$. |
| Average number of exceedances per second of some specified level, L $$\doteq e^{-L^2/2\sigma^2}$$ $$\times \frac{1}{2\pi} \left[ \frac{\int_0^\infty \omega^2 \Phi(\omega) d\omega}{\int_0^\infty \Phi(\omega) d\omega} \right]^{1/2}$$ | 41 | Stationary Gaussian random inputs (Ref. 17). | Minimum value of output or error exceedances. System order specified. | Examples studied to date, for both output and error exceedances, show that the criteria tend to select very heavily damped systems. | Originally proposed as a fatigue measure (output exceedances). |
| $p = \text{Prob} \left\{ |\epsilon| < L \right\}$ L is a specified tolerance | 54 | Stationary Gaussian random input, plus deterministic component. | Maximum value (Ref. 6, 18, and 19). System order specified (Ref. 18). | Not generally useful as formulated in Ref. 18. | Ref. 18 considered the case when the error has two components such that $E[\epsilon^2]_{t_0} = \epsilon_{s0}^2 + \sigma^2$ where $\epsilon_{s0}$ is the systematic error at time $t_0$ and $\sigma^2$ the variance of the stationary random component of the error. Then, if $$[\epsilon_{s0}] \ll \sigma, \quad \text{and} \quad L\epsilon_{s0} \ll \sigma,$$ $$\text{or} \quad \epsilon_{s0} \gg \sigma, \quad \text{and} \quad \epsilon_{s0} \doteq L$$ it is claimed that maximizing p is the same as minimizing $E[\epsilon^2]$. This is strictly true only if $\epsilon_{s0}$ is independent of the free parameters. For this case, minimizing $\sigma^2$ maximizes p independently of the above conditions. If $\epsilon_{s0}$ is not independent of the free parameters, then there is no guarantee that the above conditions hold as parameters are varied. |

TABLE I (Continued)

SUMMARY OF RANDOM INPUT PERFORMANCE MEASURES
For Linear Constant-Coefficient Systems

| MEASURE | SEE PAGE | TYPE OF SYSTEM INPUTS | ASSOCIATED CRITERION | CRITERION ASSESSMENT | REMARKS |
|---|---|---|---|---|---|
| Time-weighted mean square error (Ref. 20) $$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} w(t)\epsilon^2(t)dt$$ $w(t) > 0$ for $-\infty \leq t \leq \infty$ | 62 | Stationary random inputs. | Minimum value. System order free (Ref. 20) or specified. | Yields the same optimum system as the $\overline{\epsilon^2}$ (Wiener or Phillips) system. | For $w(t)$ statistically independent of $\epsilon(t)$ (either deterministic or nondeterministic), performance measure is shown to be equal to $(\overline{\epsilon^2})(\overline{w})$. Ref. 20 considers the case of deterministic $w(t)$ only. |
| Glover's amplitude-weighted error (Ref. 21) $$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} w(t)\epsilon^2(t)dt$$ where $w(t) = \dfrac{1}{f_d^2(t) + \delta^2}$, a nondeterministic function, $f_d(t)$ is desired signal, $\delta^2$ indicates the lowest accuracy of interest. | 75 | Stationary random inputs. | Minimum value. System order free (Ref. 21) or specified. | Tends to be very unselective when system order is specified. | For $w(t)$ not statistically independent of $\epsilon(t)$ and for Gaussian processes, the optimum system is the Wiener system multiplied by a constant (which is a function of the input and system parameters) (Ref. 21). Optimization with transfer functions of specified order (Phillips method) yields systems not related to the Phillips system in any simple manner. |
| End sigma $$S = \int_0^{\infty} p(t)F\{\epsilon(t),\ t,\ \nu_1,\ \nu_2,\ \cdots\ \nu_n\}dt$$ | 66 | Combined stationary or nonstationary random and deterministic inputs. | Minimum value (Ref. 22 and 23). | Valuable for nonstationary problems where precise formulation of $p(t)$ is possible. For stationary situations, reduces to simpler criteria, or is unsuitable. | Proposed as an all-encompassing criterion (Ref. 22). The integrand is a function of various system parameters, $\nu_1$, $\nu_2$, $\nu_3$, $\cdots$ $\nu_n$, and $p(t)$, the probability distribution of all times when the system output is utilized. Published examples all use simplified criterion given below. |
| Probabilistic square error $$\int_0^{\infty} p(t)\epsilon^2(t)dt$$ | 66 | Combined stationary or nonstationary random and deterministic inputs. | Minimum value. Maximum system order specified either explicitly or implicitly | Special form of "end sigma" criterion. Assessment as above. | Complete digital computer evaluation procedures using Legendre polynomials are available (Ref. 24, 25, 26, and 27). Analytic procedures discussed in this report. |

TABLE I (Concluded)

SUMMARY OF RANDOM INPUT PERFORMANCE MEASURES
For Linear Constant-Coefficient Systems

| MEASURE | SEE PAGE | TYPE OF SYSTEM INPUTS | ASSOCIATED CRITERION | CRITERION ASSESSMENT | REMARKS |
|---|---|---|---|---|---|
| Generalized error function (Ref. 39, 88) $$G.E.F. = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \left[ s(t - \tau) - c(t) \right]^2 dt$$ | 87 | Stationary random signal and noise. | Minimum value. System form free or specified. | When system order is free, yields Wiener system for the appropriate $\tau$ (prediction or interpolation). When the system order is specified (Phillips system), measure is difficult to minimize analytically since the expressions for G.E.F. contain exponential terms. | This measure has a compatible form for deterministic inputs $$E(\tau) = \int_{0}^{\infty} \left[ s(t - \tau) - c(t) \right]^2 dt$$ In Ref. 28, $E(\tau)$ is called ISDE (integral square delayed error) when there is zero noise. |
| Frequency-weighted mean square error (Ref. 32) $$EMS_W = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_{\epsilon\epsilon}(\omega) W(\omega) d\omega$$ | 105 | Stationary random signal and noise. | Minimum value. System form free or specified. | Free order optimum systems are related to Wiener system. Useful criterion when $W(\omega)$ can be specified so that it reflects the physical requirements of the system (see below). | $W(\omega)$ should be largest for frequency regions where error power is undesirable. $W(\omega)$ must not tend to zero as $\omega$ tends toward infinity if system order is free. |
| Mean square error derivatives (Ref. 55) $$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \sum_{n=0}^{N} a_n \left[ \frac{d^n \epsilon(t)}{dt^n} \right]^2 dt$$ | 112 | Stationary random signal and noise. | Minimum value. System order free or specified. | Simple forms of this criterion are studied in Chapter V, and it is shown that they fail to achieve good selectivity. | This measure is a form of the frequency-weighted mean square error measure $[W(\omega)$ is then a numerator polynomial in $\omega^2]$. |

Consider a pitch control system for the fighter airplane detailed in Appendix C of Ref. 1. The open-loop transfer function for the pitch loop is

$$G(s) = \left\{ \frac{4.85\left(\frac{s}{1.372} + 1\right)\left(\frac{s}{0.0098} + 1\right)}{\left[\frac{s^2}{(0.0630)^2} + \frac{2(0.0714)}{0.0630}s + 1\right]\left[\frac{s^2}{(4.27)^2} + \frac{2(0.493)}{4.27}s + 1\right]} \right\}$$

Airplane Transfer Function

$$\times \left\{ \frac{K_m K_1\left(\frac{s}{2.4} + 1\right)}{\left(\frac{s}{50}\right)^2 + \frac{2(0.7)}{50}s + 1} \right\}$$
(1)

Controller Transfer
Function

The Bode diagram for $G(j\omega)$ is shown in Fig. 2. The closed-loop system has three regions of interest defined by

(a) $|G(j\omega)| \gg 1$, over which $\left|\frac{G(j\omega)}{1 + G(j\omega)}\right| \doteq 1$

(b) $|G(j\omega)| \ll 1$, over which $\left|\frac{G(j\omega)}{1 + G(j\omega)}\right| \doteq |G(j\omega)|$

(c) $|G(j\omega)| \doteq 1$

The form of the closed-loop transfer function, $\left|\frac{G(j\omega)}{1 + G(j\omega)}\right|$, in this last region defines the "dominant modes" of the closed-loop-system dynamic response for impulse and step inputs. In most cases $G(j\omega)/[1 + G(j\omega)]$ in the region where $|G(j\omega)|$ is of the order of unity can be approximated by a first-, second-, or third-order system, the modes of which will determine the major features of the response. The open-loop amplitude asymptotes of an appropriate equivalent system for this example are shown in Fig. 2.

7
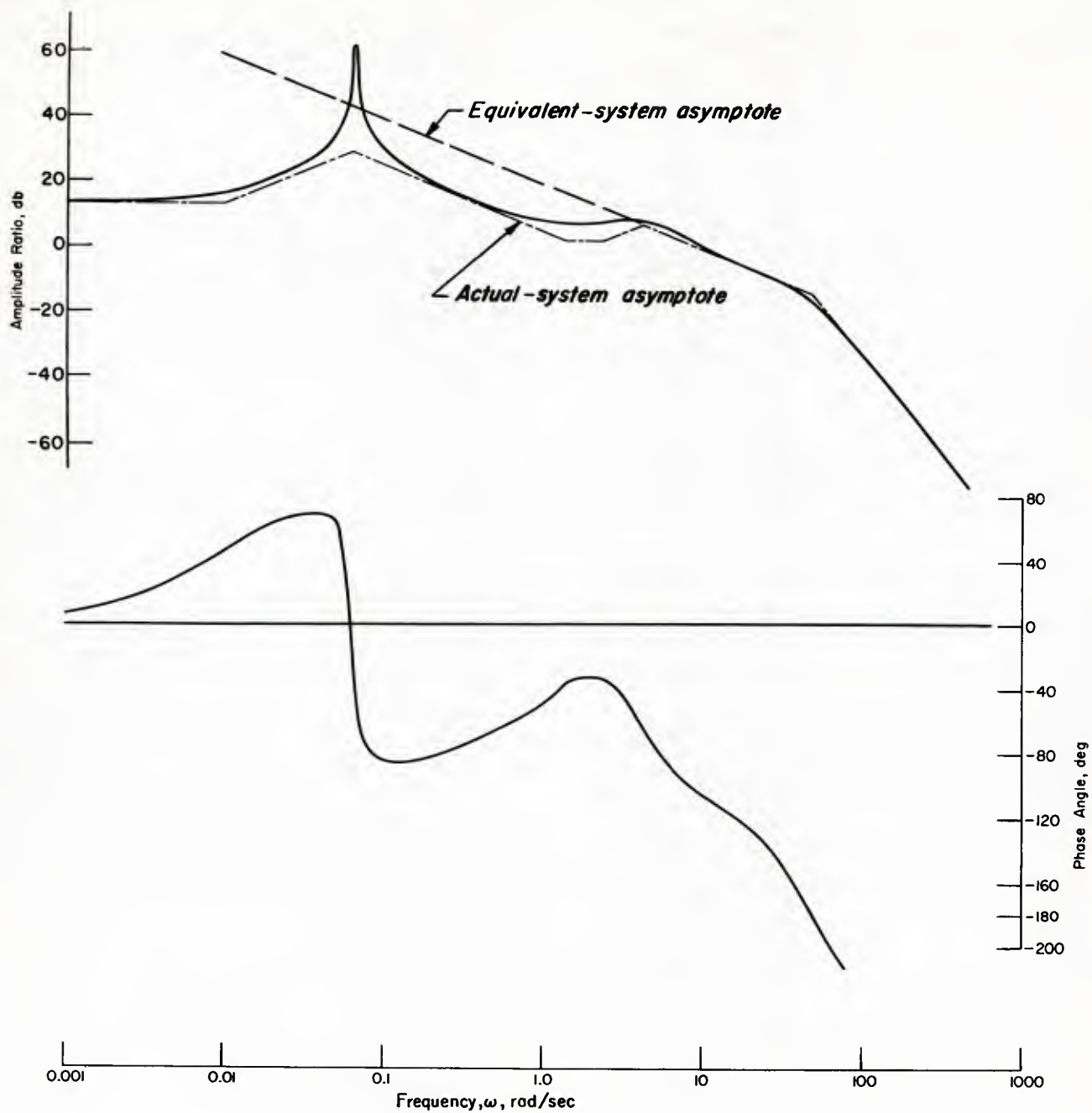
Figure 2.  Open-Loop Bode Diagram of $G(j\omega)$ Transfer Function

8

Applying this approximation to the present example yields the closed-loop $(j\omega)$ Bode diagram of Fig. 3. The Bode diagram for the exact closed-loop system is also shown for comparative purposes. It will be observed that the error of the approximation is small. If greater accuracy is required, more complicated open-loop equivalent systems can be produced by retaining more of the terms in the complete open-loop transfer function.

In this example the crossover frequency, $\omega_c$, is of the order of the servo break frequency (50 rad/sec). Usually this frequency will be $\gg \omega_c$; the effect of the associated high-frequency leads and lags can then be approximated by replacing them in either the open- or closed-loop transfer functions by a pure time delay term, $e^{-\tau s}$. A satisfactory approximation for the time delay is $\tau = -(T_{leads} - T_{lags})_{high\ frequency}$ (Ref. 1).

In general, airplane transfer function break frequencies and time constants are spaced so that $G(j\omega)$ in the region of crossover can be satisfactorily approximated by a system of not more than fourth order.

The artifice of replacing the actual flight control system by the equivalent system is extremely valuable in simplifying analysis and optimization procedures. Such simplification is usually necessary to prevent the physical realities being submerged by a flood of mathematics.

C. ASSESSMENT OF PERFORMANCE CRITERIA

The basic requirements that a criterion must satisfy if it is to be of practical use are defined in Ref. 34 as

1. validity
2. selectivity
3. ease of application

These will now be reviewed, and the concept of "compatibility" introduced. As will be shown, this concept is useful in determining the validity of random input performance criteria for flight control systems.

Validity

Validity means that the criterion yields systems that have desirable performance characteristics for the input environment of interest. Desirable characteristics would include adequate phase margin, realizable bandwidth, and, for a

9
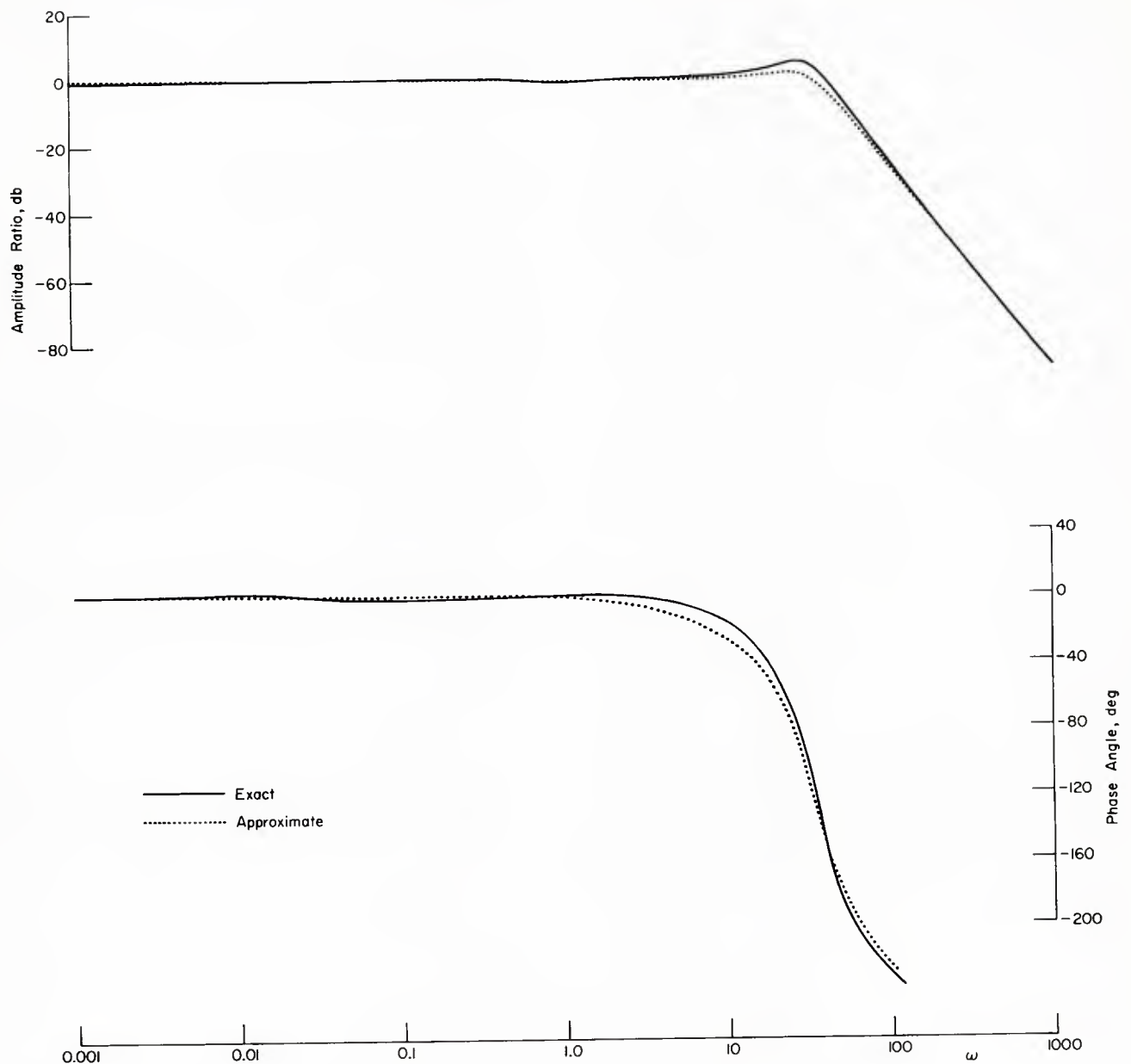
Figure 3. Comparison of Exact and Approximate Closed-Loop Bode Diagrams of $G(j\omega)/1 + G(j\omega)$ Transfer Function

10

transient input environment, good step response. Although no simple form of step response can be explicitly defined as a universally applicable optimum, it is generally accepted that fast rise time, small overshoot, and low settling time are desirable. (For full discussion of these aspects of system performance see Ref. 1.) A second-order unit numerator system having a damping ratio, $\zeta$, of about 0.7 satisfies these requirements quite well.

At first sight it might seem that an assessment of system merit based on step (or other transient) response has little relevance to the value of that system for random inputs. However, flight control systems must be satisfactory for a wide range of inputs, some (such as gusts) essentially of a random character, others (such as engine failures, pilot commands) steplike, and primarily deterministic. A random input performance criterion that yields a system having a very lightly damped step response cannot, therefore, be regarded as valid for flight control systems. Ideally, a random input performance criterion applied to a second-order unit numerator system should yield a damping ratio of approximately 0.7. This ideal provides a yardstick by which most of the criteria discussed in this report will be assessed. Although good step response is a necessary, rather than sufficient, condition for validity, criteria that fall far short of this standard can be rejected without further study, particularly since many equivalent flight control systems are only of second order.

To formalize the assessment of system merit on the basis of response to both deterministic and random inputs, the concept of "compatibility" will now be introduced. Consider a deterministic input, $r_D(t)$, and a stationary random input, $r_R(t)$. Let these inputs be applied to identical linear constant-coefficient systems having the transfer function $H(s)$, and let $M_R$ and $M_D$ be appropriate random and deterministic input performance measures, respectively. If, for any specified $r_R(t)$, an $r_D(t)$ can be found such that $M_D = M_R$, for all $H(s)$, then $M_D$ and $M_R$ are said to be <u>compatible</u> and $r_D(t)$ is called the <u>transient analog</u> of $r_R(t)$. (For compatibility only the numerical values of $M_D$ and $M_R$ need be equal; their dimensions (units) may differ.)

A brief example of a compatible criteria is given on p. 13. A more detailed discussion of compatible criteria and transient analogs for both signal and noise is given in Chapter IV. The use of transient analog inputs and compatible performance measures enables analysis of systems with random input quantities to be replaced by analysis with deterministic input quantities (such as steps and ramps),

which are much more easily visualized.  Compatibility is thus a desirable
quantity.  It does not justify criteria; the prime requirement is still validity.
However, the compatibility concept could be used to guide the search for a valid
random input performance criterion.  Instead of trying various random criteria
and then checking their (possibly unsatisfactory) validity, one could adopt the
following direct procedure:

1.  find a valid deterministic criterion
2.  find its compatible random form

Neither of these tasks may be easy.  Nevertheless, it is hoped that this system-
atic approach will be more successful than the usual procedure of suggesting
random input criteria without consideration of the deterministic response char-
acteristics of the resulting system.

Some physical significance can be attached to compatibility.  It can be
argued that the separation of flight control system inputs into "random" and
"deterministic" categories is too arbitrary.  For example, pilot inputs can
often be well approximated by step movements of the controls, the amplitude and
timing of these steps being random.  Suppose that in such an input the delay
between each successive step greatly exceeds the system settling time so that
the system error essentially settles to zero between successive steps.  This
input could be thought of

1.  as a noise-free stationary random input
2.  as a stationary sequence of steps commencing
    at time $t_0$ and continuing to $t = \infty$.

One method of optimizing the system would be to apply some standard deterministic
input performance criterion to each member of the sequence of step responses.  If
this deterministic criterion is valid (as it should be), the optimized system
step response would resemble the step response of a second-order unit numerator
system with a damping ratio of about 0.7.

Finding a valid compatible random criterion is much more difficult than it
may appear at first glance.  Many deterministic performance criteria do not have
compatible forms, and some that are compatible fall short of the ideal as regards
validity.  This last point will now be demonstrated by considering the $\overline{\epsilon^2}$ and $IE^2$
performance measures.

12

$$\overline{\epsilon^2} \;=\; \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \epsilon^2 dt \qquad \text{for stationary random inputs}$$

$$IE^2 \;=\; \int_{0}^{\infty} \epsilon^2 dt \qquad\qquad \text{for deterministic inputs commencing at } t = 0$$

It will now be shown that while these measures are compatible, their validity is less than ideal. In order to emphasize the physical significance of compatibility, only the special class of random inputs described on p. 12 will be considered. (A general proof of the compatibility of $\overline{\epsilon^2}$ and $IE^2$ for both correlated and uncorrelated signal and noise inputs is given in Chapter IV.)

Example of Compatible Criteria. It is clear that the minimum $IE^2$ criterion cannot be directly applied to stationary random quantities because $\epsilon^2$ fluctuates about a constant level and the integral would fail to converge.

Consider the special input described on p. 12, consisting of steps of random amplitude occurring at random intervals substantially exceeding the system settling time. Evaluating $IE^2$ for each of the step responses yields

$$\int_{t_0}^{t_1} \left[\epsilon(t)\right]^2 dt + \int_{t_1}^{t_2} \left[\epsilon(t)\right]^2 dt + \cdots + \int_{t_{n-1}}^{t_n} \left[\epsilon(t)\right]^2 dt$$

where   $t_0$   is the time of application of the first step

$t_1$   is the time of application of the second step

$t_{n-1}$   is the time of application of the nth step

$t_n \;=\; t_{n-1} +$ an arbitrary finite time much greater than the system settling time

If the input is stationary, this sum becomes the sum of an infinite number of integrals, and hence fails to be of use in practical calculations. In order to obtain a finite measure, the time average value of these integrals will be taken, and the performance measure modified to

$$\frac{1}{(t_n - t_{n-1}) + \cdots + (t_1 - t_0)} \left\{ \int_{t_0}^{t_1} \epsilon^2 dt + \int_{t_1}^{t_2} \epsilon^2 dt + \cdots + \int_{t_{n-1}}^{t_n} \epsilon^2 dt \right\}$$

Coalescing the limits of integration (see p. 101 of Ref. 6 and p. 136 of Ref. 8 for formal justification of this step), the criterion becomes

$$\text{minimum} \ \frac{1}{t_n - t_0} \int_{t_0}^{t_n} \epsilon^2 dt$$

For a stationary situation, $t_0 \rightarrow -\infty$ and $t_n \rightarrow +\infty$, and the criterion can be expressed in the familiar form

$$\text{minimum} \ \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} \epsilon^2 dt \ \equiv \ \text{minimum} \ \overline{\epsilon^2} \tag{2}$$

Having related $\overline{\epsilon^2}$ and $IE^2$ by elementary physical considerations, one would expect that both the minimum $\overline{\epsilon^2}$ and the minimum $IE^2$ criteria would be compatible. This can be quickly proved for a unit numerator second-order system (although the proof can be generalized for any system). The required system transfer functions are, for unity feedback,

$$H(s) \ = \ \frac{C(s)}{R(s)} \ = \ \frac{1}{s^2 + 2\zeta s + 1}, \ \frac{E(s)}{R(s)} \ = \ \frac{s^2 + 2\zeta s}{s^2 + 2\zeta s + 1} \tag{3}$$

As has been shown in Ref. 1 , for a unit step input,

$$IE^2 \ \equiv \ \int_0^{\infty} \left[ \epsilon(t) \right]^2 dt \ = \ \zeta + \frac{1}{4\zeta} \tag{4}$$

$IE^2$ is minimized by $\zeta = 0.5$. As shown on p. 141 of Ref. 8, the special random input of p. 12 could be described in conventional statistical terms by a power spectrum of the form

$$\Phi_{rr}(\omega) \ = \ \lim_{a \rightarrow 0} \frac{1}{a^2 + \omega^2} \tag{5}$$

Evaluating $\overline{\epsilon^2}$ with the above input applied to the H(s) of Eq 3 yields

$$\overline{\epsilon^2} = \lim_{a \to 0} \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \Phi_{\epsilon\epsilon}(s)ds \qquad (6)$$

$$= \lim_{a \to 0} \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \left[ \frac{(s^2 + 2\zeta s)}{(s^2 + 2\zeta s + 1)(a + s)} \right]\left[ \frac{(s^2 - 2\zeta s)}{(s^2 - 2\zeta s + 1)(a - s)} \right] ds$$

This can be evaluated using the tabulated integral forms of Appendix A, Ref. 6, whence

$$\overline{\epsilon^2} = \zeta + \frac{1}{4\zeta} = IE^2 \qquad (7)$$

Thus the $\overline{\epsilon^2}$ and $IE^2$ criteria are compatible. A general proof (applicable to all stationary random inputs) showing that these performance measures are compatible for any linear system is given in Chapter IV. The value of this compatibility is rather diminished by the fact that minimum $IE^2$ is generally regarded as only moderately valid because the $\zeta = 0.5$ step response has a relatively high over-shoot and a long settling time compared to, say, the optimum ITAE second-order unit numerator system ($\zeta = 0.76$).

It is interesting to examine the difficulties in constructing a compatible criterion for ITAE. With the same input as above, applying the ITAE criterion to each of the step responses yields the criterion

$$\text{minimum} \int_{t_0}^{t_1} t|\epsilon|dt + \text{minimum} \int_{t_1}^{t_2} t|\epsilon|dt + \cdots + \text{minimum} \int_{t_{n-1}}^{t_n} t|\epsilon|dt$$

Taking the time average value of these integrals, the criterion may be modified to

$$\text{minimum} \frac{1}{t_n - t_0}\left\{ \int_{t_0}^{t_1} t|\epsilon|dt + \int_{t_1}^{t_2} t|\epsilon|dt + \cdots + \int_{t_{n-1}}^{t_n} t|\epsilon|dt \right\}$$

There appears to be no standard random form to which this expression can be con-densed, i.e., no random criterion compatible with ITAE has been found. Of course,

15

this is not a proof that no compatible criterion exists; if one could be found, it would be extremely valuable. Further discussion of compatible criteria is given in Chapter IV.

## Other Requirements for Performance Criteria

"Validity" has been discussed at some length; the remaining requirements of "selectivity" and "ease of application" can be dealt with more briefly.

"Selectivity" implies sharp differentiation between "good" systems and those which are merely "acceptable." Of the performance measures examined in Ref. 1, ITAE is particularly satisfactory in this respect. For a step input to a second-order unit numerator system, the minimum ITAE is 1.96 at $\zeta = 0.76$, rising to 2.25 at $\zeta = 0.61$ and $\zeta = 0.90$. Since the ITAE criterion maintains good selectivity for high-order systems and has been favorably received in almost all relevant references (e.g., Ref. 31 and 35), it will be taken as setting an acceptable standard for selectivity.

"Ease of application" demands that the criterion should be readily expressible in terms of system parameters, and that convenient procedures for its evaluation should exist.

CRITERIA DEFINED SOLELY AS FUNCTIONS OF ERROR

This chapter is concerned with performance measures expressible in the form $\overline{f(\epsilon)}$, the sole independent variable being $\epsilon$, the error between the actual system output and a specified desired output. The majority of the criteria that have been proposed for linear constant-coefficient systems with random inputs are of this form. In particular, the criterion of minimum mean square error has been extensively studied for stationary random inputs. Outstanding contributions to the study of this criterion have been made by Wiener (Ref. 4) and Phillips (Ref. 3), and throughout this report reference will frequently be made to Wiener optimization and Phillips-type optimization. Therefore, for ease of reference, the Wiener and Phillips-type optimization procedures are briefly summarized below. The essential difference between the procedures is that in Phillips-type optimization the form of the system (the order of the transfer function numerator and denominator) is prescribed by the analyst, whereas the form of the Wiener system depends only upon the spectra describing the input environment and desired output.

## A. BRIEF SUMMARY OF WIENER OPTIMIZATION

Wiener investigated the problem of recovering a stationary random signal, $f(t)$, from an input comprised of signal plus noise, $n(t)$, when the desired output, $c_d(t)$, is equal to the signal advanced or retarded by $\eta$ sec. The cases where $\eta > 0$, $\eta = 0$, and $\eta < 0$ are referred to as prediction, smoothing, and interpolation (or lagging), respectively. No restrictions are imposed other than that the optimum system should be linear and physically realizable, i.e., no output can arise without a prior input. A block diagram representation of this problem is given in Fig. 4. Components are indicated by their transfer functions, for brevity.
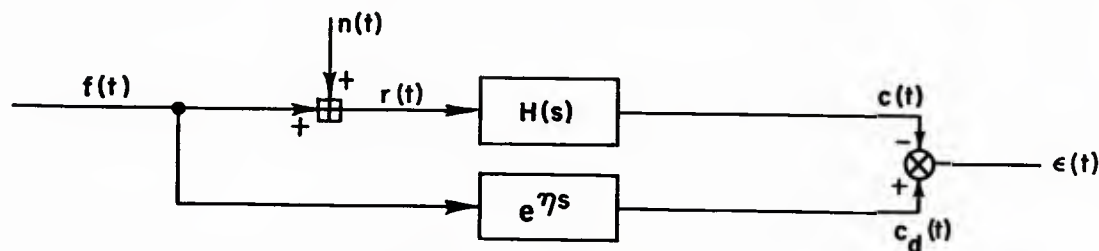


Figure 4. Block Diagram for Wiener Optimization

H(s) is the transfer function of the system under consideration, and its output, $c(t)$, is denoted by $c_W(t)$ when $H(s)$ becomes $H_W(s)$, the Wiener (optimum linear) system. Using the ergodic hypothesis, the time average square error equals the ensemble average square error

$$\overline{\epsilon^2} = E(c_d - c)^2 \tag{8}$$

where $\epsilon(t) = c_d(t) - c(t)$, the difference between desired and actual outputs, and E denotes the ensemble average

Let $H_W(s) = \mathcal{L}\left[h_W(t)\right]$ denote the physically realizable linear constant coefficient system which minimizes $\overline{\epsilon^2}$. $H_W(s)$ is found by considering any other realizable linear system, $h_W(t) + ag(t)$; and imposing the condition that for all a (a parameter) and any $g(t)$, this system yields a larger $\overline{\epsilon^2}$ than that due to $h_W(t)$; i.e.,

$$E(c_d - c_W - ac_g)^2 - E(c_d - c_W)^2 \geq 0 \tag{9}$$

where $c_g(t)$ is the additional output due to $g(t)$, and $c_W(t)$ is the output of the Wiener system

Expanding Eq 9 gives

$$E(c_d - c_W)^2 - 2aE\left[(c_d - c_W)c_g\right] + a^2E(c_g^2) - E(c_d - c_W)^2 \geq 0 \tag{10}$$

$$\therefore \quad -2aE\left[(c_d - c_W)c_g\right] + a^2E(c_g^2) \geq 0 \tag{11}$$

Because the second term is always positive, the inequality can hold for all a only if

$$E\left[(c_d - c_W)c_g\right] = 0 \tag{12}$$

18

This is the necessary and sufficient condition for $c_W$ to be the output of the linear system that yields minimum $\overline{\epsilon^2}$. Using the convolution relationship for physically realizable linear systems,

$$c_g(t) = \int_0^\infty g(\tau) r(t - \tau) d\tau \qquad (13)$$

Equation 12 can be written

$$aE\left[ \int_0^\infty g(\tau) r(t - \tau) \left\{ c_d(t) - c_W(t) \right\} d\tau \right] = 0 \qquad (14)$$

Interchanging integrations,

$$a\int_0^\infty g(\tau) \left\{ R_{rc_d}(\tau) - R_{rc_W}(\tau) \right\} d\tau = 0 \qquad (15)$$

where the crosscorrelation is defined as

$$R_{ab}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} a(t) b(t + \tau) dt = E[a(t)b(t + \tau)]$$

where a and b are general stationary random signals

Equation 15 will be true for all $g(\tau)$, $\tau \geq 0$, and all a, only if the expression in braces is equal to zero, i.e.,

$$R_{rc_d}(\tau) = R_{rc_W}(\tau) \qquad \tau \geq 0 \qquad (16)$$

Usually the desired output is either the signal or a prediction of this signal $\eta$ sec in the future:

19

$$c_d(t) = f(t + \eta) \tag{17}$$

Because $c_W(t) = \int_0^\infty h_W(u)r(t - u)du$, Eq 16 becomes the familiar Wiener-Hopf integral equation:

$$R_{rf}(\tau + \eta) = \int_0^\infty h_W(u)R_{rr}(\tau - u)du \qquad \tau \geq 0 \tag{18}$$

The solution of Eq 18 yields the optimum linear system. Alternatively, Eq 18 can be expressed in the frequency domain as (Ref. 10, p. 92)

$$H_W(j\omega) = \frac{1}{\Phi_{rr}^+(j\omega)}\left[\frac{e^{j\omega\eta}\Phi_{rf}(j\omega)}{\Phi_{rr}^-(j\omega)}\right]_+ \tag{19}$$

where $\Phi_{rf}(j\omega)$ = the cross-spectrum between the desired output and the sum of signal and noise

$\Phi_{rr}^+(j\omega)$ = a factor of $\Phi_{rr}(\omega)$ containing all the left-half-s-plane poles and zeros of $\Phi_{rr}(\omega)$

$\Phi_{rr}(\omega)$ = $\Phi_{rr}^+(j\omega)\Phi_{rr}^-(j\omega)$ (the input power spectrum)

$[\ ]_+$ means expand in partial fractions, and then keep only terms with left-half-plane poles

Equations 18 and 19 constitute the principal results of the Wiener optimization theory.

B. BRIEF SUMMARY OF PHILLIPS-TYPE OPTIMIZATION

The method of optimization developed by Phillips differs from that described above in that the form of $H_p(s)$, the Phillips optimum system transfer function, is prescribed. The error power spectrum is calculated for general values of the numerator and denominator coefficients of $H_p(s)$. The mean square error is then obtainable from the following relationship:

$$\overline{\epsilon^2} \;=\; \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_{\epsilon\epsilon}(\omega)\,d\omega \tag{20}$$

The evaluation of this integral is facilitated by the use of integral tables
given by Phillips (Ref. 3 and in a more convenient form in Ref. 1, 6, and 36).
The mean square error is thus obtained in literal terms involving the coefficients
of $H_p(s)$, and the resulting expression is then minimized with respect to these
coefficients.

To illustrate a typical Phillips-type optimization, a second-order zero-
position-error system will be considered; its transfer function is

$$H(s) \;=\; \frac{1}{s^2 + 2\zeta s + 1}$$

The signal power spectrum is $\Phi_{ff}(s) = \dfrac{1}{-s^2 + a^2}$, the noise is zero, and the desired
output is equal to the signal. The resulting mean square error has been graphed
in Fig. 5. For a specified value of a, the Phillips filter has the $\zeta$ indicated
by the dotted line. In this example, the mean square error criterion is unselec-
tive in that the off-optimum mean square error is very little greater than the
minimum. However, discussion of the merit of $\overline{\epsilon^2}$ as a performance criterion
relative to other criteria will be delayed until Chapter III, after demonstrating
that many apparent alternative criteria are in fact equivalent to minimum mean
square error.

C. DETERMINATION OF THE MINIMUM MEAN SQUARE ERROR SYSTEM
   WITHOUT RESTRICTION OF LINEARITY OR CONSTANCY OF COEFFICIENTS

When the signal and noise are stationary random processes, the Wiener system,
by definition, yields a lower mean square error than any other physically realiz-
able linear constant-coefficient system. It is logical to inquire whether a
time-varying and/or nonlinear physically realizable system could yield a smaller
mean square error in such an input environment. Strictly, the answer to this
question lies beyond the scope of this report, which is primarily concerned with
time-invariant linear systems; nonlinear systems are usually studied separately
because their analysis demands specialized techniques, and because the generality

$$H(s) = \frac{1}{s^2 + 2\zeta s + 1} \quad ; \quad \Phi_{rr}(\omega) = \frac{1}{a^2 + \omega^2}$$

$$\overline{\epsilon^2} = \frac{1 + 4\zeta^2 + 2\zeta a}{4\zeta(1 + 2\zeta a + a^2)}$$

a=0.0
a=0.1
a=0.25
a=0.4
a=0.5
a=0.6
a=0.8
a=1.0
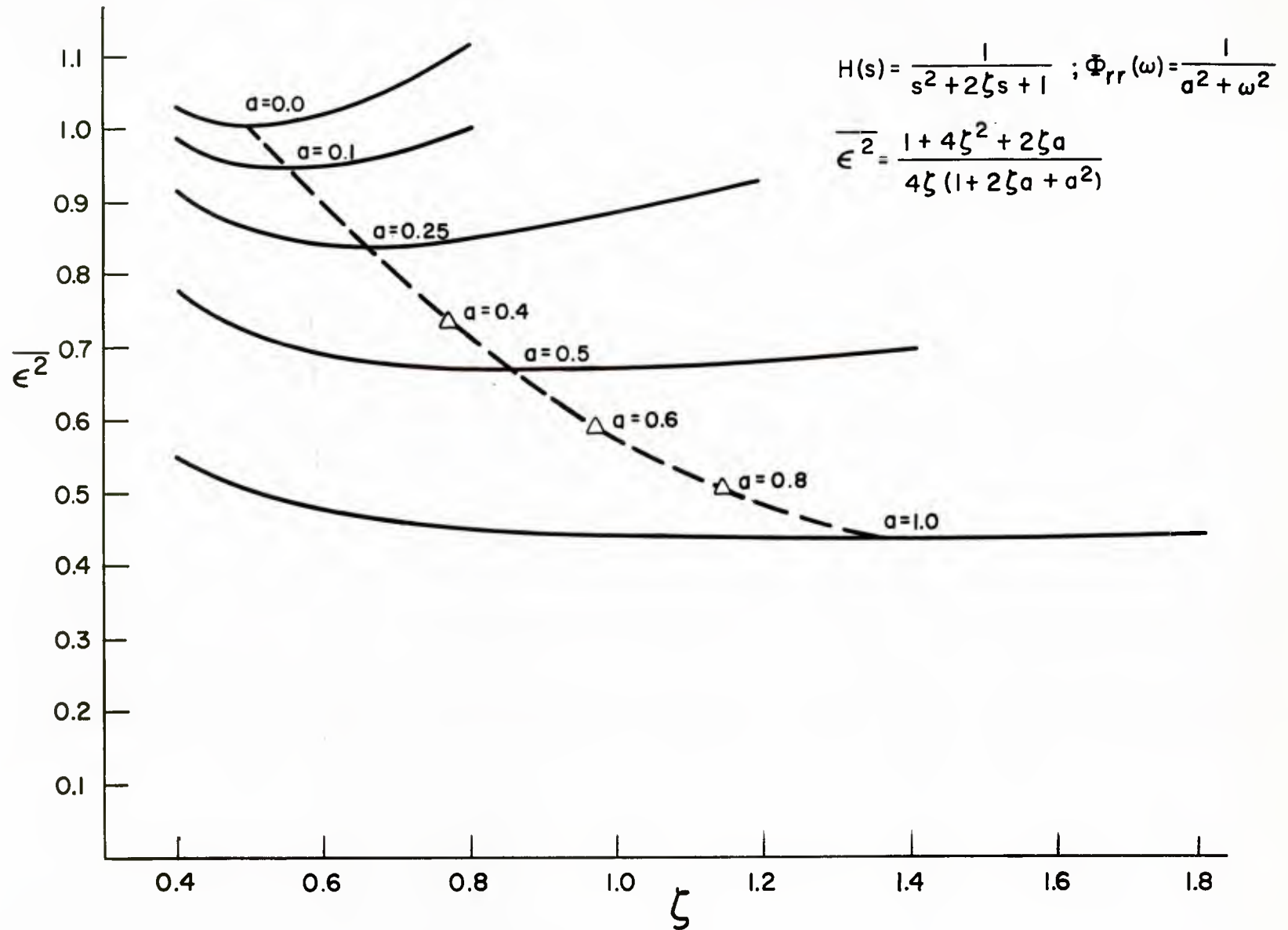
Figure 5.   Effect of Input Break Frequency on Mean Square Error of a Second-Order Unit Numerator System

of the results obtained is limited. (For example, the response of a nonlinear system to a deterministic input may change radically with a small change in the input amplitude.) However, for stationary random inputs, it is sometimes possible to obtain quite useful and general results without imposing the restriction of linearity. In particular, it can be shown that with a stationary random Gaussian input environment, the Wiener system yields a lower mean square error than any other physically realizable system, linear or nonlinear, with or without time-varying coefficients.

In order to demonstrate this result, a brief departure from purely linear analysis is necessary. This is amply justified by the importance of the result obtained.

## Formulation of the Ensemble Mean Square Error

It is necessary first to consider a general (i.e., not necessarily Gaussian) input environment, and to derive the formula for the minimum mean square error system. Sherman (Ref. 13) quoted this formula, but did not prove it. In fact, no proof was found in any of the references consulted by the authors of this report. The derivation given below was obtained by utilizing some results given by Cramer (Ref. 37).

Consider the system illustrated in Fig. 4 (p. 17) where now the sole restriction upon the output, c, is that it be the output of a physically realizable system; i.e., c depends only upon the past and present values of r. The system is permitted to be linear or nonlinear and time-varying or time-invariant. The ensemble mean square error is

$$E(c_d - c)^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (c_d - c)^2 p(c_d, c) dc_d dc \qquad (21)$$

where $p(c_d, c)$ is the joint proba-
bility density function of $c_d$ and c

So that Cramer's results can be used without modification, it will be assumed that the system operates only upon sampled values of $r(t)$ available only at a finite number, n, of sampling instants prior to t, the present time. For this condition, $c(t)$ will be some function, $\psi$, of the present and past input samples; i.e.,

$$c(t) = \psi(r_0, r_1, r_2, \cdots, r_n) \qquad (22)$$

$$\text{where} \quad r_0 = r(t)$$
$$r_1 = r(t - T)$$
$$r_2 = r(t - 2T)$$
$$\cdots \cdots \cdots$$
$$r_n = r(t - nT)$$

T is the sampling period

Making use of Eq 22 the ensemble mean square error, Eq 21, can be written as

$$E(c_d - c)^2 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (c_d - c)^2 p(c_d, r_0, r_1 \cdots r_n) dc_d dr_0 dr_1 \cdots dr_n \quad (23)$$

where $p(c_d, r_0, r_1, \cdots, r_n)$ is the joint probability density function of $c_d$ and all the past sampled values of $r$ that contribute to $c$

The joint probability density function $p(c_d, r_0, r_1, \cdots, r_n)$ can be rewritten as

$$p(c_d, \hat{r}) = p(c_d/\hat{r})p(\hat{r}) \qquad (24)$$

where $\hat{r}$ denotes the sequence of past sampled input values, $r_0, r_1, r_2, \cdots, r_n$

$p(c_d/\hat{r})$ is the conditional probability density function defined the probability density function of $c_d$, assuming that the particular input sequence $\hat{r}$ has occurred.

Inserting Eq 24 into Eq 23, and rearranging terms, yields

$$E(c_d - c)^2 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dr_0 dr_1 \cdots dr_n p(\hat{r}) \int_{-\infty}^{\infty} (c_d - c)^2 p(c_d/\hat{r}) dc_d \quad (25)$$

24

## Minimization of the Ensemble Mean Square Error

The minimization of $E(c_d - c)^2$ follows the procedure in Ref. 37, p. 271-272. The following points about Eq 25 should be noted:

1. c occurs only in the inner integral over $c_d$

2. because c is a function of $\hat{r}$ alone, it is constant when integrating over $c_d$

3. the integral over $c_d$ is the second moment of $c_d$ taken about c, given that a general $\hat{r}$ has occurred.

Equation 25 will be minimized if c is chosen such that the integral over $c_d$ is minimized at <u>each</u> value of $\hat{r}$ that can possibly arise from the given signal and noise. The well-known fact that the second moment is minimized when taken about the mean implies that c must be equal to the conditional mean of $c_d$ given $\hat{r}$, i.e.,

$$c = E(c_d/\hat{r}) = \int_{-\infty}^{\infty} c_d p(c_d/\hat{r}) dc_d \tag{26}$$

This is the general formula for the absolute optimum mean square error system.

<u>Interpretation of Eq 26 for General Inputs</u>. The meaning of Eq 26 can best be understood by considering the case where $c_d$ equals the signal; as usual, r is the sum of signal and noise. The present and past sampled values of r, relative to the present time, t, are available as inputs to the absolute optimum system. From these observations on the sum of signal and noise, the best estimate of the signal alone is the ensemble average value of the signal at time, t, utilizing the actual sequence, $\hat{r}$, that has occurred. The formula for the absolute optimum system (Eq 26) is intuitively satisfying because it makes use of all the information that observations on the sum of signal and noise have conveyed about the signal alone; i.e., $p(c_d/\hat{r})$. Unfortunately, the determination of the conditional mean is in general very difficult if the signals involved are non-Gaussian.

<u>Interpretation of Eq 26 for Gaussian Signal and Noise</u>. When the signal and noise are Gaussian, both $c_d$ and r are Gaussian. In particular, the sequence $\hat{r} = r_0, r_1 \cdots r_n$ (defined in Eq 24) has a multivariate Gaussian joint probability function. Cramer (Ref. 37, p. 314, 315) has shown that conditional density

functions of Gaussian signals are Gaussian. In addition, he has shown that the conditional mean is a <u>linear</u> function of the given variables $\hat{r} = r_0, r_1, \cdots r_n$. It is now permissible to allow n, the number of sampled values, to become arbitrarily large and the sampling period to become as small as may be desired. Neither of these limiting processes will affect the linearity of the absolute optimum system. Therefore, the absolute optimum system is linear, and if the Gaussian signal and noise are also stationary then, by definition, this system must be the Wiener system.

## D. SUMMARY OF RESULTS ON MEAN SQUARE ERROR CRITERIA

The results obtained on the Wiener, Phillips, and absolute minimum mean square error systems are summarized on the "family tree" of Fig. 6, which shows how these systems are related. It will be demonstrated later in this report that many other criteria yield related systems, and that many of these latter systems are simply minimum mean square error systems. In particular, family trees for criteria expressible as $f(\epsilon)$ and for time-weighted criteria are given on pp. 28 and 63, respectively. (It would be possible to join appropriate branches from each of these trees. However, the resulting single tree would have many branches, resulting in a fairly complicated presentation which would tend to obscure the simplicity of many of the relationships involved.)

## E. BENEDICT AND SONDHI'S PERFORMANCE MEASURE (Applicable to Stationary Gaussian Processes Only, Both Wiener and Phillips-Type Optimization)

For stationary Gaussian processes, Benedict and Sondhi (Ref. 38) have shown that the linear system (Wiener or Phillips-type) which minimizes $\overline{\epsilon^2}$ also minimizes any performance measure of the form

$$\overline{f(|\epsilon|)} = \overline{\sum_k |\epsilon|^{n_k}} = \sum_k \overline{|\epsilon|^{n_k}} \tag{27}$$

where $n_k$ is positive (but
not necessarily an integer)
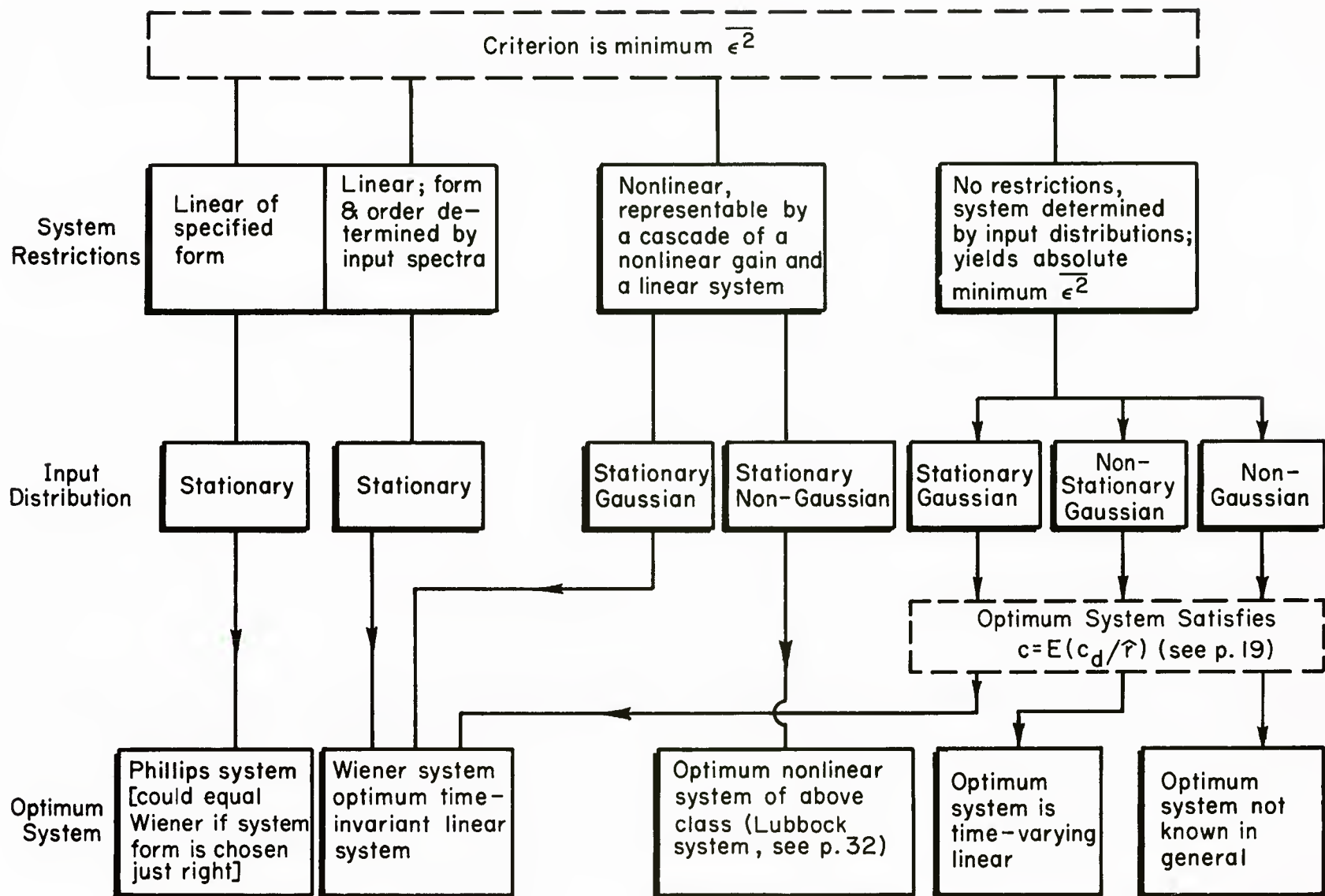
Figure 6. Summary of Minimum Mean Square Error Systems
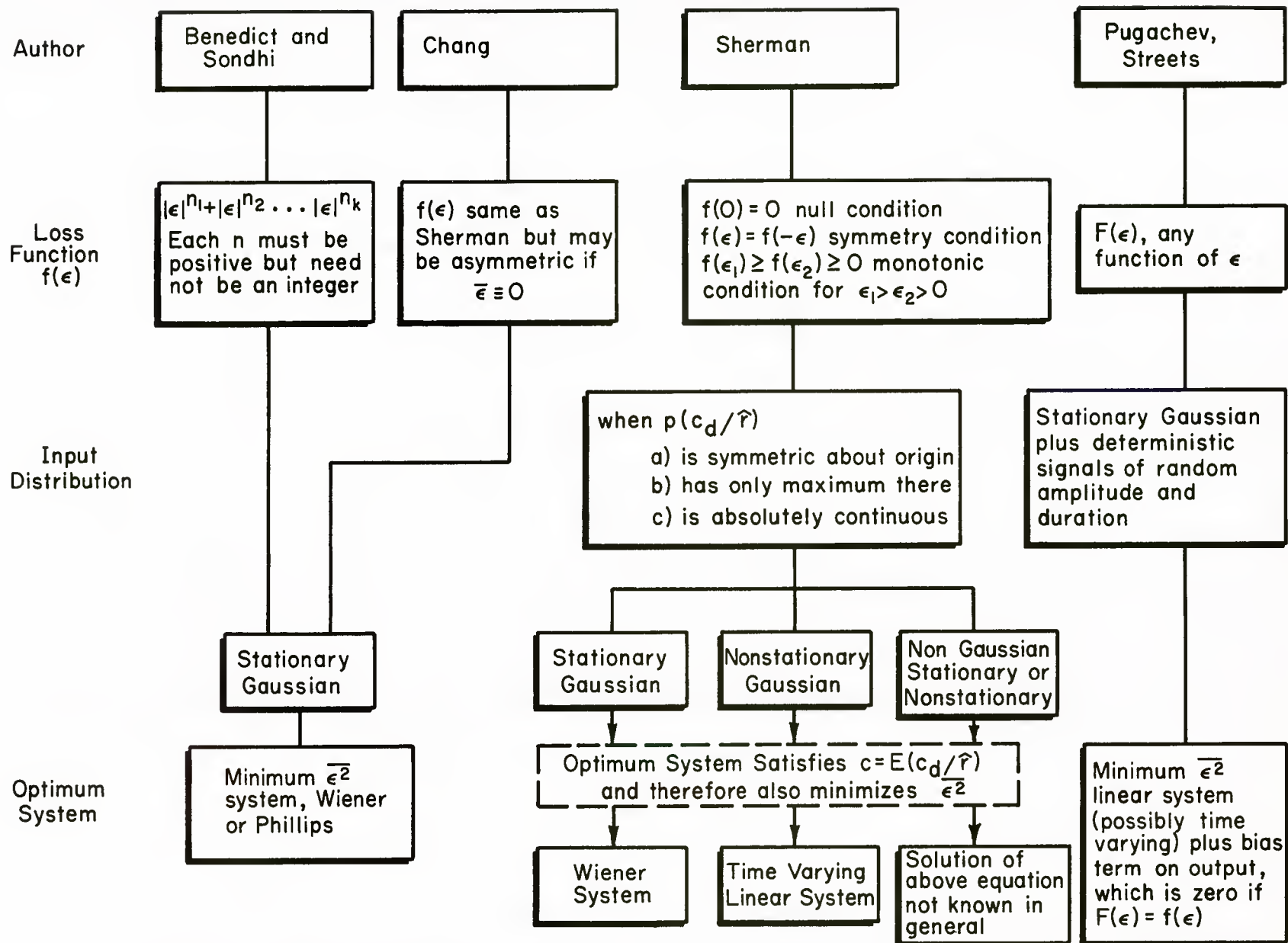
Figure 7. Summary of Minimum $\overline{f(\epsilon)}$ Systems

<u>Proof</u>

Consider a general example

$$\overline{f(|\epsilon|)} = \overline{|\epsilon|^{n_1}} + \overline{|\epsilon|^{n_2}} + \overline{|\epsilon|^{n_3}} + \cdots \qquad (28)$$

where $n_1$, $n_2$, $n_3$ $\cdots$ are real and positive integers or nonintegers (rational or nonrational) and the number of $n_k$'s used is arbitrary

Each term $\overline{|\epsilon|^{n_k}}$ can be evaluated using the ensemble average, which for Gaussian processes is

$$\overline{|\epsilon|^{n_k}} = E\left[|\epsilon|^{n_k}\right] = 2\left(\frac{1}{\sigma\sqrt{2\pi}}\int_0^\infty \epsilon^{n_k} e^{-(\epsilon^2/2\sigma^2)} d\epsilon\right) \qquad (29)$$

Using the integral tables of Ref. 39, p. 201,

$$\overline{|\epsilon|^{n_k}} = \frac{2^{n_k/2}}{\sqrt{\pi}}\left[\Gamma\left(\frac{n_k+1}{2}\right)\right]\sigma^{n_k} \qquad (30)$$

where $\quad \sigma = \sqrt{\overline{\epsilon^2}}$ as usual

$\Gamma(x) =$ the Gamma function

Substituting Eq 30 into Eq 28 yields

$$\overline{f(|\epsilon|)} = \sum_k \frac{2^{n_k/2}}{\sqrt{\pi}}\Gamma\left(\frac{n_k+1}{2}\right)\sigma^{n_k} \qquad (31)$$

From Eq 31 it is apparent that $\overline{f(|\epsilon|)}$ can be minimized only by minimizing $\sigma$. Therefore, the system that minimizes $\overline{\epsilon^2}$ minimizes all possible forms of $\overline{f(|\epsilon|)}$ as defined in Eq 27 and 28.

29

It can be shown that $f(|\epsilon|)$ is a nondecreasing function of $\epsilon$. The slope is

$$\frac{df}{d\epsilon} = \sum_k n_k \epsilon^{n_k - 1} \tag{32}$$

which for $\epsilon > 0$ is always positive and nonzero. It would appear that one could construct a great variety of $f(|\epsilon|)$'s by appropriate choice of the k's and $n_k$'s. A typical $f(|\epsilon|)$ is sketched below.
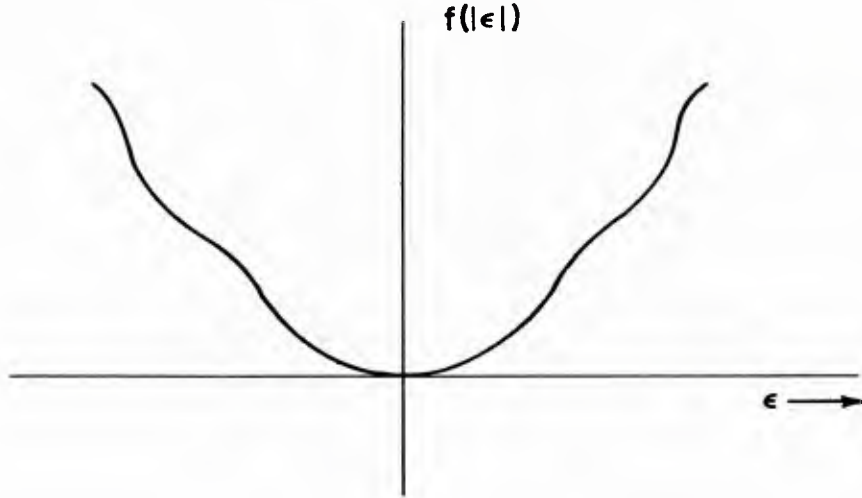


Figure 8. Typical Nondecreasing $f(\epsilon)$ Loss Function

Note that $f(|\epsilon|)$ could be generalized to the form

$$\overline{f(|\epsilon|)} = \sum_k a_{n_k} \overline{|\epsilon|^{n_k}} \tag{33}$$

without changing the fact that $f(|\epsilon|)$ is minimized by minimizing $\overline{\epsilon^2}$.

30

All the above results could have been derived using the analyses of Sherman and Chang presented later in this chapter. The original derivation of Benedict and Sondhi was given because it is particularly easy to follow and forms a useful preparation for the discussion of Sherman's results that now follows.

## F. SHERMAN'S PERFORMANCE MEASURE

Benedict and Sondhi (Ref. 38) considered a fairly general $f(\epsilon)$ with stationary Gaussian random signals and found that $\overline{f(\epsilon)}$ was minimized by minimizing $\overline{\epsilon^2}$. The question naturally arises as to whether or not a similar simplification occurs when the signal and noise are nonstationary and/or non-Gaussian random processes. This problem has been solved by Sherman (Ref. 13), who showed that minimizing $E[f(\epsilon)]$, defined below, is accomplished by minimizing $E[\epsilon^2]$. (Ensemble averages are used since the ergodic hypothesis will not hold if the input environ- is nonstationary.) Many other performance measures can be expressed as special cases of Sherman's measure. These special cases are detailed below, and their interrelationships are summarized on the "family tree" on p. 28.

Sherman's "loss function," $f(\epsilon)$, has the following properties:

$$
\left.
\begin{aligned}
&f(0) \;=\; 0 &&\text{null property} \\
&f(\epsilon) \;=\; f(-\epsilon) &&\text{symmetry property} \\
&\left.\begin{aligned} f(\epsilon_2) \geq f(\epsilon_1) \geq 0 \\ \text{for}\quad \epsilon_2 \geq \epsilon_1 \geq 0 \end{aligned}\right\} \text{monotonic property}
\end{aligned}
\right\} \tag{34}
$$

This $f(\epsilon)$ is similar to the loss function of Benedict and Sondhi (Ref. 38) (illustrated in Fig. 3 ) in that $f(\epsilon)$ must never decrease for increasing $|\epsilon|$, but it differs in that it may have zero slopes at some points or in some regions.

The optimum system should minimize the ensemble average loss, which is

$$
E[f(\epsilon)] \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(c_d - c)p(c_d,\, c)\,dc_d\,dc \tag{35}
$$

By a line of reasoning similar to that in Section C, Eq 35 can be written in a form similar to Eq 26. In the derivation of Eq 36, one merely substitutes $f(\epsilon)$

31

for $\epsilon^2$, and then the equation corresponding to Eq 25 for $E\left[f(\epsilon)\right]$ becomes

$$E\left[f(c_d - c)\right] \; = \; \int_\infty^\infty \cdots \int_{-\infty}^\infty p(\hat{r})dr_0 dr_1 \cdots dr_n \int_{-\infty}^\infty f(c_d - c)p(c_d/\hat{r})dc_d \qquad (36)$$

The system that minimizes Eq 36 depends on the particular conditional density function, $p(c_d/\hat{r})$, that results from the given signal and noise distribution. Now, if the conditional probability density function, $p(c_d/\hat{r})$, satisfies the following sufficient conditions (derived from some more general conditions stated by Sherman, Ref. 13),

1. symmetric about the origin of $c_d$
2. has its only maximum there
3. is continuous

then a lemma in probability theory quoted by Sherman states that $E\left[f(c_d - c)/\hat{r}\right]$ is minimized when c equals the conditional mean of $c_d$ given $\hat{r}$, i.e.,

$$c \; = \; E\left[c_d/\hat{r}\right] \qquad (37)$$

$$\text{where} \quad c \; = \; c(t)$$
$$c_d \; = \; c_d(t)$$
$$\hat{r} \; = \; \text{the sequence } r_0, \; r_1, \; r_2 \cdots r_n$$
$$\text{as defined in Section C, Eq 24}$$

Equation 37 describes the system that yields the absolute minimum $E\left[f(\epsilon)\right]$ at the present time, t. This is the same system that minimizes the $\overline{\epsilon^2}$, which, as noted in Section C, may well be nonlinear and/or time-varying, depending on the statistics of the signal and noise present. Kalman (Ref. 40) remarks that as far as he is aware, it is not known what is the most general random signal and noise for which the conditional distribution function satisfies conditions 1, 2, and 3 above.

Special Cases

$\underline{f(\epsilon) = \epsilon^2}$. It was shown in Section C that when $f(\epsilon) = \epsilon^2$, Eq 37 for the optimum system minimizes $E(\epsilon^2)$ without the requirement that $p(c_d/\hat{r})$ satisfy conditions 1, 2, and 3 (following Eq 36). This was also noted by Kalman in Ref. 40.

Gaussian Signal and Noise. In the important case of nonstationary Gaussian signal and noise, the conditional distribution of $c_d$ given $\hat{r}$ is Gaussian (as shown by Cramer, Ref. 37, p. 315), which satisfies conditions 1, 2, and 3. As demonstrated in Section C, the output of the optimum system (Eq 26) is a linear function of the past values of the sum of signal and noise. If the signal and noise are stationary as well as Gaussian, the absolute optimum system for all $f(\epsilon)$ as defined in Eq 26 is the Wiener system.

Chang's Extension of Sherman's Results for Gaussian Processes. In many situations it is desirable that the mean error, $\overline{\epsilon}$, be equal to zero. For these cases Chang (Ref. 10) has shown that, for Gaussian processes, minimizing $\sigma_\epsilon^2$ also minimizes Sherman's $\overline{f(\epsilon)}$. This will be true even if the symmetry property of $f(\epsilon)$ in Eq 34 is relaxed. The above result is almost intuitively obvious, since a Gaussian process is determined by its mean and variance, and with zero mean only the variance can be reduced to reduce $\overline{f(\epsilon)}$.

Kaufman's (Ref. 16) Performance Measure. Kaufman has used as a performance measure

$$\chi = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \sum_{n=1}^{N} c_n \epsilon^{2n} dt = \sum_{n=1}^{N} c_n \overline{\epsilon^{2n}} \tag{38}$$

for cases where $\overline{\epsilon^2}$ is not sufficiently meaningful. He restricted his analysis to the following situations:

1. signal and noise are statistically independent processes with non-Gaussian probability density functions which are symmetric about the mean

2. only linear systems are considered.

Kaufman showed how to evaluate and minimize $\chi$ analytically. The procedure was approximate and quite involved, and only calculations of $\overline{\epsilon^2}$ and $\overline{\epsilon^4}$ were demonstrated.

The system that yields the absolute minimum $\chi$ depends on the conditional density function of the signal, given the past history of the sum of signal and noise (as shown by Eq 37). The information given in Item 1 above is not sufficient to apply Sherman's simplification to this performance measure. But if the conditional density function of $c_d$ given $\hat{r}$ satisfies Sherman's conditions, then $f(\epsilon)$ may as well be replaced by $\epsilon^2$. The absolute minimum $\overline{\epsilon^2}$ system quite probably will not be linear since the signal and noise are non-Gaussian.

33

## G. PUGACHEV'S AND STREETS' MINIMIZATION OF GENERAL F($\epsilon$) CRITERIA WHEN THE SIGNAL IS PARTIALLY DETERMINISTIC

In this section, optimization procedures for a class of inputs more general than that hitherto considered will be summarized. These inputs are allowed to have partially deterministic components. This corresponds to an important group of flight control inputs where the input may have a component of known amplitude and form (as in bang-bang control), but where the timing between successive applications of this input component is random. As described in Section E of this chapter, the minimum $\overline{\epsilon^2}$ system also minimizes a wide class of other $\overline{F(\epsilon)}$ measures. Pugachev (Ref. 14) has shown that, when the input contains Gaussian signals plus a partially deterministic component, the minimum $\overline{\epsilon^2}$ system is either equal to, or very simply related to, the minimum $\overline{F(\epsilon)}$ system, where $F(\epsilon)$ is a general loss function. Pugachev's results will now be briefly summarized and a brief account given of their application by Streets (Ref. 41). This will be followed by a summary of Lubbock's procedure (Ref. 42) for optimizing a class of nonlinear systems to yield minimum $\overline{\epsilon^2}$ with non-Gaussian inputs. The reader concerned only with random Gaussian inputs will find little use for the analyses of Pugachev, Streets, and Lubbock, and he is advised to skip these sections and turn to the Summary on page 39.

Pugachev (Ref. 14) considers an input signal given by

$$r(t) = s(t) + n(t) + \sum_{k=1}^{N} \alpha_k f_k(t) \tag{39}$$

where  $s(t)$ = stationary Gaussian signal

$n(t)$ = stationary Gaussian noise

$\alpha_k f_k(t)$ = additional signal terms; the $f_k(t)$ are known functions of time, and the $\alpha_k(t)$ are random variables with a known Gaussian joint probability density function, $p(\alpha_1 \alpha_2 \cdots \alpha_N)$

It is also assumed that $s(t)$ and $n(t)$ both have zero means, known auto- and cross-correlation functions, and are statistically independent of the additional signal terms. The desired output, $c_d(t)$, is a linear function of the total signal, $s(t) + \sum \alpha_k f_k(t)$.

34

Pugachev shows that for an arbitrary error criterion, $F(\epsilon)$ [not restricted to be a nondecreasing even function of $(\epsilon)$], the optimum system is the $\overline{\epsilon^2}$ optimum linear system plus a constant bias term, k, as shown in Fig. 9.
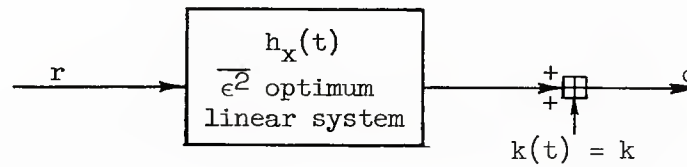


Figure 9.  Pugachev's Optimum System

The error has a mean value of k and changing the form of $F(\epsilon)$ changes k, but does not affect $h_x(t)$, the $\overline{\epsilon^2}$ optimum system weighting function. If $F(\epsilon) = f(\epsilon)$, an even function  of the class defined by Sherman, then k is zero.

Streets (Ref. 41) has worked out two problems of the type considered by Pugachev.  He demonstrated that a theoretical (linear, time-varying) system can be designed which, by taking advantage of the partially deterministic character of the input, yields a lower $\overline{\epsilon^2}$ than a Wiener system.  The problems studied have Gaussian white noise for $n(t)$, while $s(t)$, the purely random signal component, is zero in Eq 39.  Streets described the partially deterministic signal as follows.  The signal is a stationary process which consists of a sequence of random steps [random ramps in the second problem] of constant-amplitude [velocity] segments (see Fig. 10 ).  The amplitude changes at "event points" that are Poisson distributed in time.  The amplitudes have a Gaussian distribution with zero mean. The amplitudes before and after an event point are independent.  The desired output equals the signal.
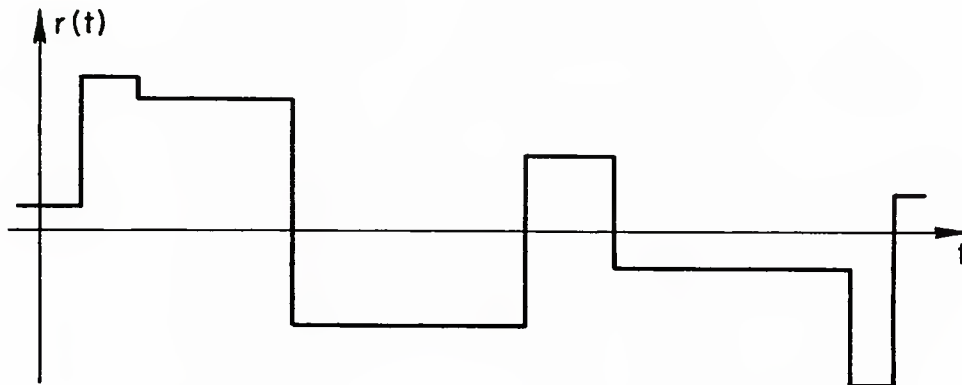


Figure 10.  Partially Deterministic Signal

35

The signal described above has a power spectrum of the form $\frac{1}{\omega^2 + v^2}$ (Ref. 50, p. 237). If the noise has a white spectrum $\left[\Phi_{nn}(\omega) = \text{constant}\right]$, the Wiener system is a first-order filter. In competition with this, Streets' theoretical optimum system requires the following three components:

1. a perfect event-point detector

2. an adaptive device which monitors the time from the last event point and adjusts the system parameters accordingly. At each event point this device sets the output signal equal to the expected mean and other output initial conditions equal zero

3. a linear, time-varying filter plus a bias term when using non-even criteria.

For low-frequency signal-to-noise ratios of 1 and 100, the improvement in performance given by Streets' theoretical system is such that the mean square errors are reduced to 90 percent and 55 percent (respectively) of the Wiener values. Similar results were obtained for the second problem using random ramps.

Streets remarks that the improvement in performance given by the theoretical system over the Wiener system is not as great as might be expected. This is because the input signal in each example is a "near Gaussian" signal. A more non-Gaussian signal would be expected to offer a larger potential improvement through use of a nonlinear or adaptive system.

It is important to note the following points about Streets' theoretical system:

1. it is optimum for the criterion of Pugachev

2. no other adaptive or nonlinear system can give a smaller $\overline{\epsilon^2}$

3. it can be analytically designed and is physically realizable except for the perfect event-point detector.

Streets assumed a perfect event-point detector to simplify the analysis and to delineate the absolute lower bound on the mean square error. If a Wiener system is close to this lower bound, there is little need to attempt to approximate the Streets theoretical optimum system, since any physically realizable event-point detector must operate in a noisy environment and may miss event points or have a delay in detecting them. Streets states that both these conditions will seriously degrade the performance of the time-varying system, but of course will not affect the performance of the Wiener system. Streets also states that if the idealized signal is a poor model of the actual signal, the performance of

the theoretical system is degraded to a much larger extent than that of the Wiener system, since the Wiener system does not depend on the detection of event points.

## H. LUBBOCK'S PROCEDURE FOR SYNTHESIS OF A CLASS OF NONLINEAR SYSTEMS HAVING MINIMUM $\epsilon^2$

Lubbock (Ref. 42) showed how to minimize the mean square error for a special class of nonlinear systems subjected to stationary random inputs with a general (i.e., not necessarily Gaussian) distribution. The class of nonlinear systems is that which can be represented by a nonlinear gain operating upon the input followed by a linear system (see Fig. 11).

$$r(t) \longrightarrow \boxed{f(r)} \longrightarrow \boxed{h(\tau)} \longrightarrow c(t) \qquad (40)$$
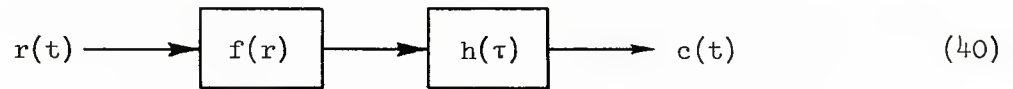
Figure 11. Nonlinear Gain Followed by a Linear System

Lubbock's procedure is directly of interest in the context of this report as an example of how a fixed-form nonlinear system may be optimized for minimum mean square error. This could be advantageous for flight control systems subject to markedly non-Gaussian inputs.

The input and output are related by

$$c(t) \;=\; \int_0^\infty f\big[r(t - \tau)\big] h(\tau) d\tau \qquad (41)$$

This is a special case of the following equation, which defines a class of filters, $\eta_1$:

$$c(t) \;=\; \int_0^\infty K\big[r(t - \tau),\; \tau\big] d\tau \qquad (42)$$

Lubbock assumes that the kernel function, $K(r, \tau)$, can be expanded in terms of a set of orthogonal functions, $\theta_a(x)$ (which are polynomials in x of degree a). The output then becomes

37

$$c(t) \; = \; \sum_{a=0}^{\infty} \int_{0}^{\infty} \theta_a \left[ r(t - \tau) \right] h_r(\tau) d\tau \qquad\qquad (43)$$

which can be synthesized by a multipath filter as shown in Fig. 12.
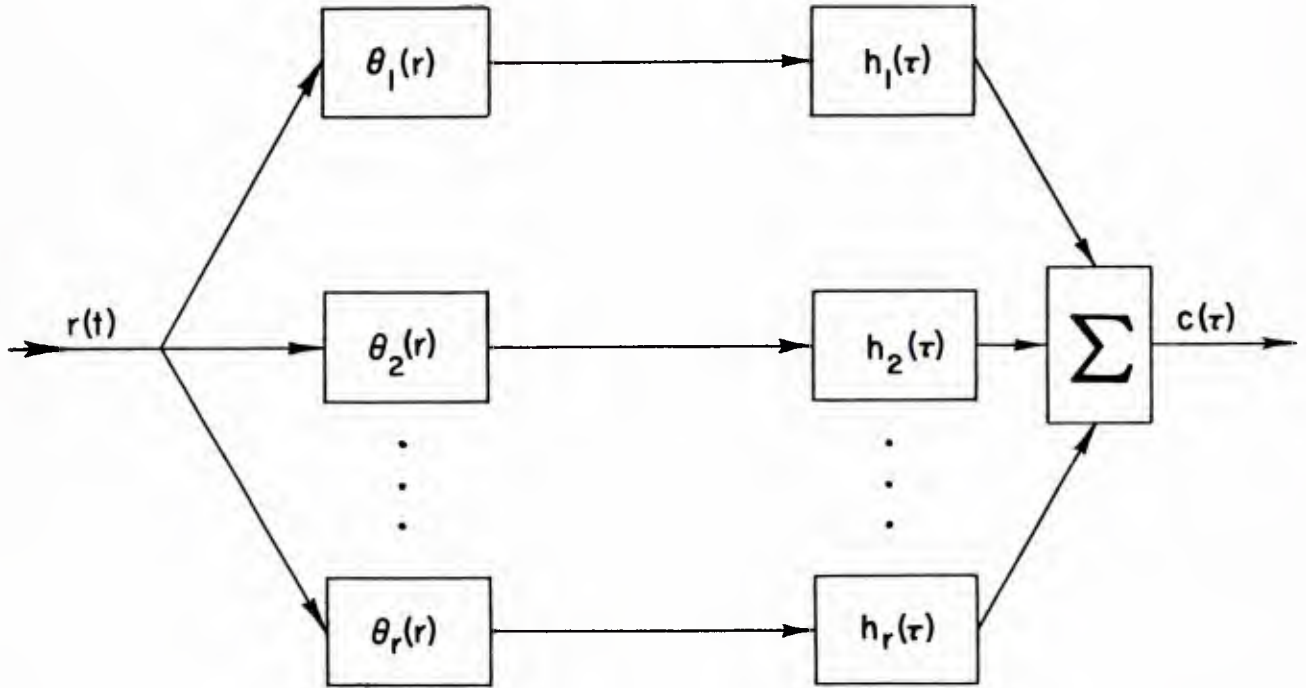


Figure 12.   Lubbock's Optimum Multipath System

Lubbock claims that by using a finite number of terms in the expansion, the filter can be physically realized and made to approximate any system defined by Eq 42 with any desired accuracy (assuming that the expansion converges).

An integral equation for the optimum filter of class $\eta_1$ is derived by variational calculus.  The information required to solve this equation is the joint probability density of the input, $p_2(r_1, r_2, \tau_1 - \tau_2)$, and the joint probability density, $p_2(r_1, c_d, \tau_1)$, between the input and the desired output.  Lubbock proceeds to formally expand each density function in a double series involving the first probability density functions, $p(r)$ and $p(c_d)$; polynomials, $\theta_a(r)$,

which are orthonormal with respect to $p(r)$ and $p(c_d)$; and the cross-correlation between the outputs of the polynomials (see Fig. 10). The restrictions on the joint probability density function required to validate this procedure are not enumerated by Lubbock, nor are they known to the present authors. Using the truncated series representation for the joint probability density functions, the integral equation for the optimum system becomes a finite number of simultaneous integral equations.

Lubbock notes that there are situations in which the optimum general system of class $\eta_1$ reduces to the Wiener system. The conditions for this simplification to take place define a class of joint probability density functions which includes Gaussian density functions.

## I.  SUMMARY

This chapter has been concerned with random input performance measures that are expressible as functions of error alone, the error being defined as the difference between the actual output and some specified or desired output. The most important result presented has been the demonstration that for a very wide class of such $f(\epsilon)$ loss functions, $\overline{f(\epsilon)}$ is minimized by the Wiener or Phillips system, provided that the input is Gaussian. Many flight control inputs have approximately Gaussian distributions (e.g., atmospheric turbulence, see Ref. 43 and 44). Hence, for many situations of practical interest in flight control optimization, the optimum system is the same for minimum $|\overline{\epsilon}|$, $\overline{\epsilon^2}$, $\overline{\epsilon^4}$, etc., and changing the loss function accomplishes no real change in the optimum system. Since it is so difficult to "escape" from the minimum $\overline{\epsilon^2}$ criterion, it is worth reiterating its advantages and deficiencies. As noted in the Introduction, the principal requirements for a performance criterion for flight control systems are validity, selectivity, and ease of application. The minimum $\overline{\epsilon^2}$ criterion may be briefly assessed in terms of these qualities as follows:

<u>Validity</u>. Minimum $\overline{\epsilon^2}$ is only moderately valid as a performance criterion, because it frequently selects damping ratios somewhat lower than would be considered optimum on the basis of transient response. Figure 5 illustrates this point: for an integrated white noise input, the second-order system considered gave minimum $\overline{\epsilon^2}$ at $\zeta = 0.5$.

Selectivity.  In all the cases that have been examined in the course of preparing this report, the optimum $\overline{\epsilon^2}$ system yielded a mean square error only a little less than that given by a wide range of off-optimum systems.  For the unit numerator second-order system of Fig. 5, varying $\zeta$ between 0.4 and 0.6 raised $\overline{\epsilon^2}$ only 2.5 percent above the minimum.  A similar lack of selectivity is exhibited by $IE^2$ for transient inputs (see Fig. 16 of Ref. 1).  (By comparison, the degree of selectivity displayed by the ITAE performance measure for step inputs is generally regarded as acceptable in this respect, and this gives a 7.5 percent increase in the measure for a change of $\zeta$ of ±0.1 from the optimum of $\zeta = 0.76$ for the above system.)

Ease of Application.  The present study is directed principally at "pencil and paper" methods of optimization, and it is essential that any criterion selected possess analytic forms which are simple enough for manual calculations. It is fortunate in this regard that (as noted in the Introduction) flight control systems can almost always be represented by low-order "equivalent systems."  This simplification keeps Phillips-type optimization within the bounds of practicality. For high-order systems and input spectra, the tabulated Phillips integral forms for $\overline{\epsilon^2}$ can become very lengthy.  However, none of the examples in this report required more than a few man-hours for the evaluation of these integral forms. The $\overline{\epsilon^2}$ criterion is thus judged to be adequately easy to apply to linear systems.

# CHAPTER II

## EXCEEDANCE CRITERIA AND CRITERIA FOR MIXED RANDOM AND DETERMINISTIC INPUTS

Section A of this chapter discusses exceedance criteria, i.e., criteria defined in terms of a certain fixed level of the error or output. Usually one attempts to minimize the average number of times per second that the specified level is crossed. Alternatively, one may seek to minimize the <u>probability</u> that the error will exceed the specified level; such a "probability criterion" was proposed by Zadeh and Ragazzini in Ref. 18 and is examined in Section B of this chapter. Zadeh and Ragazzini considered inputs comprised of a deterministic signal plus random noise. It will be shown that for such inputs the probability criterion is of little value, although mean square error criteria can be successfully applied. A discussion of this last point is given in Section C.

## A. EXCEEDANCES

Thorson and Bohne (Ref. 45), among others (Ref. 46 and 47), have discussed methods of calculating the expected total number of exceedances of a prescribed level experienced by a specified vehicle during a given mission. This number of exceedances, coupled with information on the fatigue characteristics of the structure, can be used to determine the expected life of the vehicle. Total exceedance numbers differ from the performance measures previously discussed in that they are not primarily intended to assess dynamic performance, but instead describe another fundamental parameter of system effectiveness, i.e., the vehicle's expected life. However, it is logical to inquire whether exceedance concepts can be used to form dynamic performance measures. In order to answer this question, some specific formulas for exceedances are required, and these will now be presented.

Exceedances can be expressed in two forms (which are frequently equivalent):

1.  the average number of exceedances of a given level by the output (or error quantity) per unit time (or per single mission for fatigue calculations to determine the total permissible number of missions)

2.  the fraction of the total time during which the output (or error quantity) exceeds the given level.

For stationary random inputs, the ergodic hypothesis permits form 2 to be expressed
as 

    3. the probability that the output or error quantity will
       exceed the given level.

In Ref. 18 this last form is proposed for the optimization of systems subjected
to inputs having both random and deterministic components. This probability
criterion is examined in Section B of this chapter, and is shown to be difficult
to apply in nontrivial cases. Throughout the section that now follows, station-
ary random inputs will be assumed.

## Formulas for Average Exceedances Per Second

Rice (Ref. 17) has shown that for a Gaussian signal, $v(t)$, with zero mean,
the expected number of axis crossings per second (including crossings with both
positive and negative slopes) is

$$N_0 = \frac{1}{\pi}\left\{ \frac{-\left[\frac{d^2 R_{vv}(\tau)}{d\tau^2}\right]_{\tau=0}}{R_{vv}(0)} \right\}^{1/2} = 2\left[ \frac{\int_0^\infty f^2 \Phi_{vv}(f)\,df}{\int_0^\infty \Phi_{vv}(f)\,df} \right]^{1/2} = \frac{1}{\pi}\left[ \frac{\int_{-\infty}^\infty \omega^2 \Phi_{vv}(\omega)\,d\omega}{\int_{-\infty}^\infty \Phi_{vv}(\omega)\,d\omega} \right]^{1/2} \quad (44)$$

where $f = \frac{\omega}{2\pi}$ = frequency

For rectangular spectra ranging from $\omega_a$ to $\omega_b$, the expected number of axis
crossings per second is

$$\frac{1}{\pi}\left[ \frac{1}{3}\frac{(\omega_b^3 - \omega_a^3)}{(\omega_b - \omega_a)} \right]^{1/2}$$

When $\omega_a = 0$ this reduces to $\frac{1.16}{2\pi}\omega_b$. Hence for many spectra of practical interest
the number of zero crossings, $N_0$, can serve as an indication of bandwidth. $N_0$
times the total operating time may also be employed in a fatigue criterion, since
it indicates the total expected number of stress reversals. In situations where
knowledge of $N_0$ is significant, the average number of maxima per second, $N_M$, may
also be important. This is given by Rice as:

$$N_M = \frac{1}{2\pi} \left\{ \frac{\left[ -\dfrac{d^4 R_{VV}(\tau)}{d\tau^4} \right]_{\tau=0}}{\left[ \dfrac{d^2 R(\tau)}{d\tau^2} \right]_{\tau=0}} \right\}^{1/2} = \frac{1}{2\pi} \left[ \frac{\int_0^\infty \omega^4 \Phi_{VV}(\omega)\,d\omega}{\int_0^\infty \omega^2 \Phi_{VV}(\omega)\,d\omega} \right]^{1/2} \tag{45}$$

For rectangular spectra of the type considered in the previous paragraph, this reduces to

$$N_M = \frac{1}{2\pi} \left[ \frac{3}{5} \frac{\left( \omega_b^5 - \omega_a^5 \right)}{\left( \omega_b^3 - \omega_a^3 \right)} \right]^{1/2} \tag{46}$$

When $\omega_a = 0$, the average number of maxima per second is simply $\dfrac{0.775}{2\pi}\, \omega_b$.

Of much more direct interest as a fatigue criterion and a possible performance measure is the average number of exceedances per second of some specified output level, L. For Gaussian inputs this is given by Rice (Ref. 17) as

$$N_L = \frac{e^{-L^2/2\sigma^2}}{2} \times N_0 \tag{47}$$

where $N_0$ = average number
of zero crossings per second

Strictly, this formula only applies when $L \gg \sigma$. The accuracy of the formula for smaller values of L has been investigated by Press, Meadows, and Hadlock in Ref. 43. They conclude that for $L/\sigma > 2$, the formula is in most cases valid. For values of $L/\sigma < 2$, it tends to slightly underestimate the number of exceedances. It is noted in Ref. 43 that for moderately flexible airplanes, the formula is 10 to 15 percent low at $L/\sigma = 1$ and 2 to 3 percent low at $L/\sigma = 2$. This would certainly be acceptable for purposes of systems optimization; hence Rice's formulas will be retained in the discussion that now follows.

Exceedance Numbers as Performance Criteria

The conventional use of exceedances as fatigue criteria relates to the number of _output_ exceedances above a given level. To form dynamic performance criteria

43

it would seem to be more appropriate to consider the number of _error_ exceedances. However, there is some interest in determining whether conventional output exceedance fatigue measures can also be used as performance measures. Therefore, in this section both output and error exceedances will be investigated for a second-order unit numerator system subjected to a stationary Gaussian input, and the variation of these exceedance numbers with $\zeta$ will be examined. The principal results obtained are

1.  minimizing the output (or error) exceedances is _not_ equivalent to minimizing the mean square output (or error)

2.  the direct application of error exceedance formulas to some input environments is hampered by the failure of certain integrals to converge. Apart from this, the criterion is fairly easy to apply, and is valid and selective when the low-frequency ratio of signal to noise power spectra is fairl  high (50:1). Unfortunately, for lower values of this ratio (5:1) the number of error exceedances tends to be minimized by excessively high $\zeta$ for the second-order system investigated

3.  further investigation of error exceedance criteria is recommended to overcome the convergence problem and to explore a suggestion (Ref. 57) that they are used by human pilots performing tracking tasks.

Output Exceedances

An example will now be given to illustrate point 1 above. The system considered has the transfer function

$$H(s) \;=\; \frac{1}{s^2 + 2\zeta s + 1} \tag{48}$$

The input power spectrum is

$$\Phi_{ss}(\omega) \;=\; \Phi_{rr}(\omega) \;=\; \frac{1}{a^2 + \omega^2} \tag{49}$$

and zero noise is assumed. Combining Eq 48 and 49, the output power spectrum is

$$\Phi_{cc}(\omega) \;=\; \left[ \frac{1}{(s^2 + 2\zeta s + 1)(s + a)} \right]\left[ \frac{1}{(s^2 - 2\zeta s + 1)(-s + a)} \right]_{s=j\omega} \tag{50}$$

44

In order to evaluate $N_0$, the number of axis crossings, it is necessary to compute $\int_0^\infty \omega^2 \Phi(\omega) d\omega$, and $\int_0^\infty \Phi(\omega) d\omega$ in Eq 44. These integrals may be evaluated by use of the Phillips integral formulas. The final result is

$$N_0 = \frac{a}{\pi(2\zeta + a)} \tag{51}$$

The average number of exceedances per second of an output level L is

$$N_L = \frac{a}{2\pi(2\zeta + a)} \cdot e^{\frac{-L^2 2a(1 + 2a\zeta + \zeta^2)}{(2\zeta + a)}} \tag{52}$$

Figures 13 and 14 illustrate how the number of exceedances varies with $\zeta$ for a = 0.1, and for L = 1, 3, 5, and 7. As would be expected, the exceedances decrease monotonically with increasing $\zeta$; this is physically reasonable since increasing $\zeta$ implies that the system is becoming more sluggish and less liable to overshoot.

To obtain a physical "feel" for the meaning of this result, consider Fig. 15. This figure illustrates the variation with $\zeta$ of $\sigma^2_{output}$, the mean square output. ($\sigma^2_{output}$ was evaluated by computing $\int_{-j\infty}^{j\infty} \Phi_{cc}(s) ds$ using Phillips integrals.) Comparing Fig. 15 with Fig. 13 and 14, it is seen that the fall-off of $N_L$ with increasing $\zeta$ is more rapid than the fall-off of $\sigma^2_{output}$. This is made clear in Fig. 16 which shows the ratio $N_L/\sigma$ plotted versus $\zeta$ for L = 1, 3, 5, and 7. For each of these values of L, increasing $\zeta$ causes $N_L$ to diminish more rapidly than $\sigma$ as shown by the monotonic decrease of the ratio $N_L/\sigma$. Figures 13, 14, and 15 show that minimizing either $\sigma^2_{output}$ or $N_L$ yields $\zeta \rightarrow \infty$. Minimization of these output quantities thus appears unpromising as a performance criterion and will not be pursued further in this report.

### Error Exceedances

The unit numerator second-order system will now be used to investigate error exceedances as performance measures. Consider the system illustrated in Fig. 17
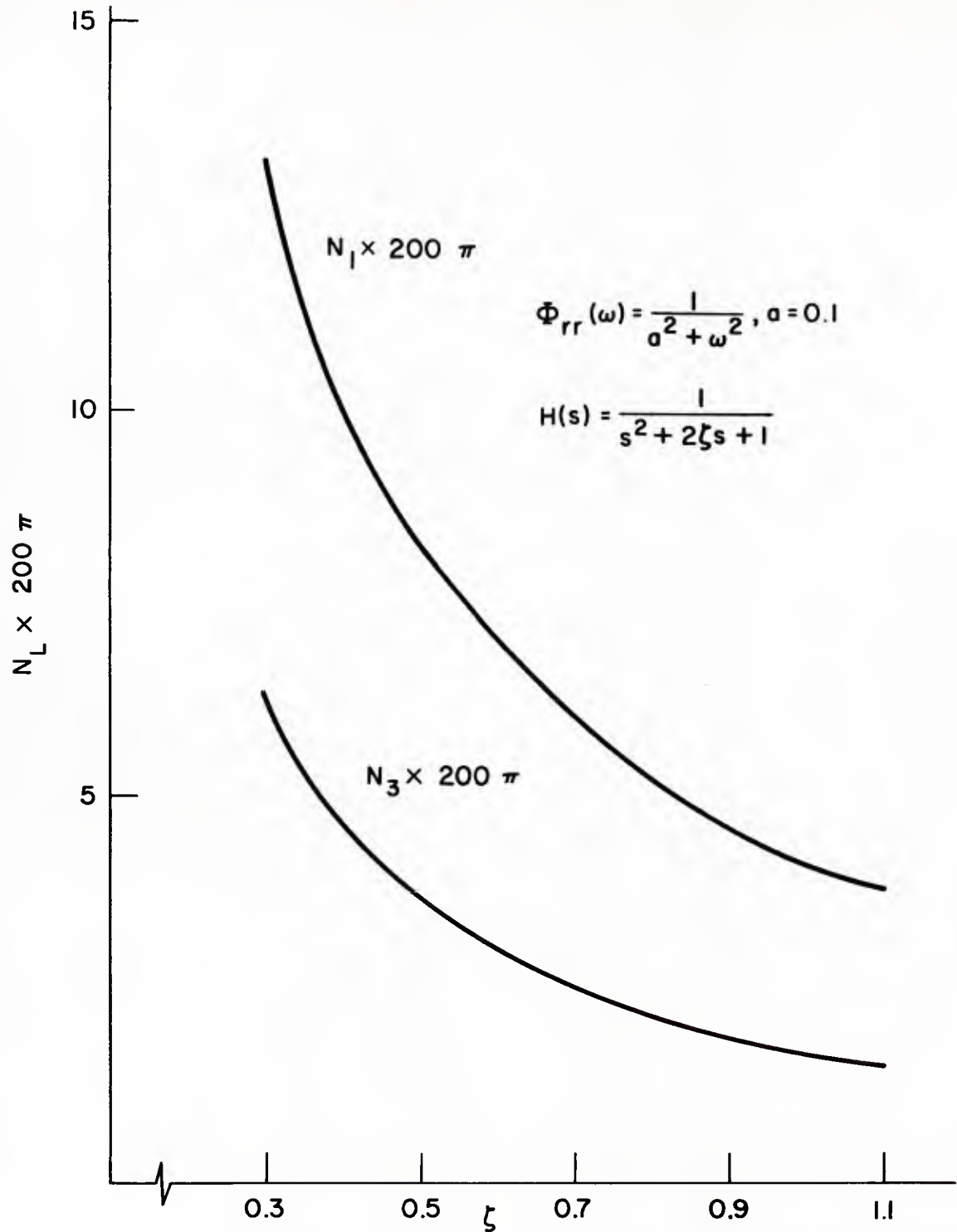
Figure 13.  Output Exceedances $N_1$ and $N_3$ vs $\zeta$ for a Second-Order Unit Numerator System
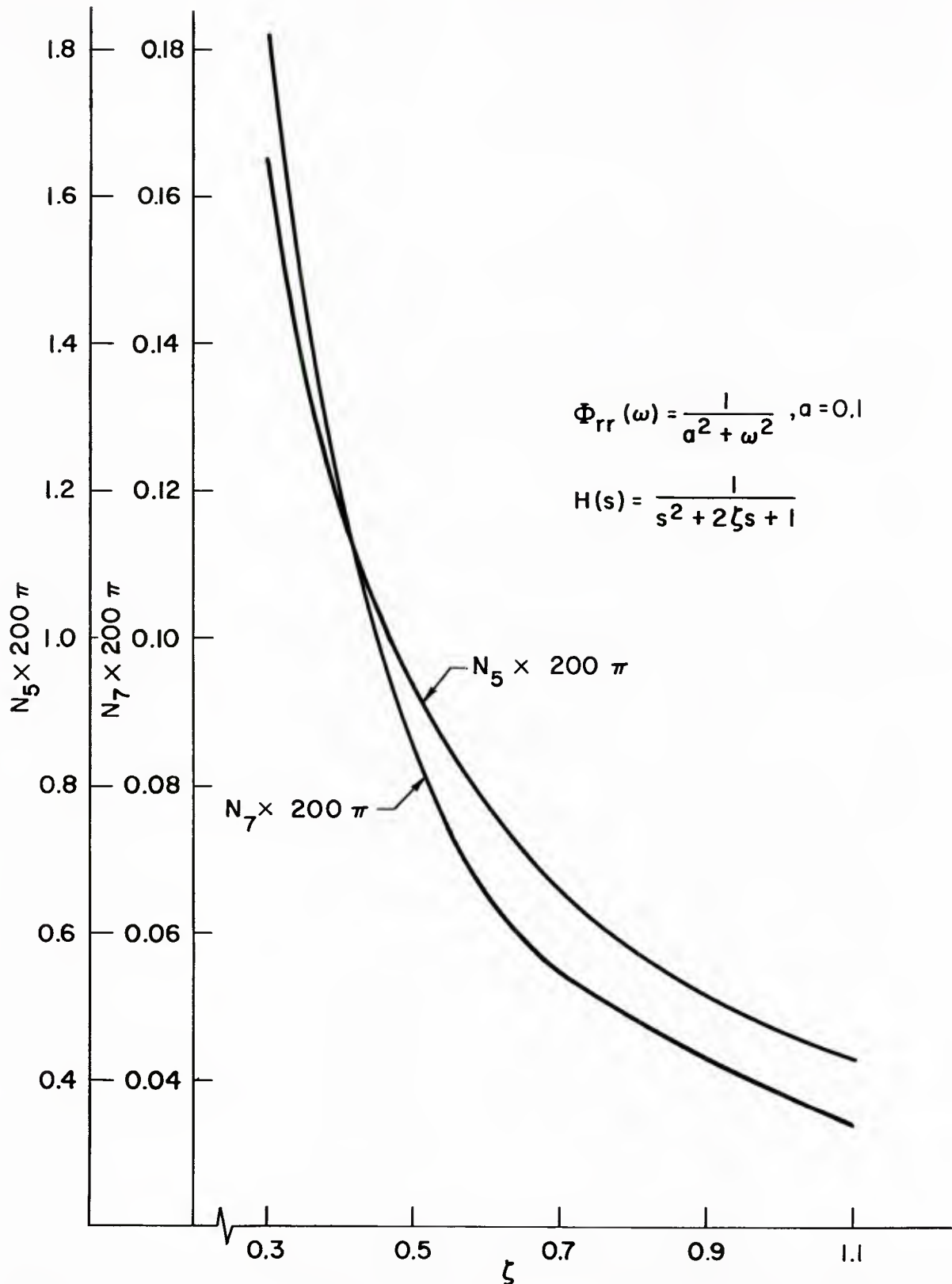
Figure 14. Output Exceedances $N_5$ and $N_7$ vs $\zeta$ for a Second-Order Unit Numerator System

$$\Phi_{rr}(\omega) = \frac{1}{a^2 + \omega^2} \;, a = 0.1$$

$$H(s) = \frac{1}{s^2 + 2\zeta s + 1}$$

47

$$\Phi_{rr}(\omega) = \frac{1}{a^2 + \omega^2} , \quad a = 0.1$$

$$H(s) = \frac{1}{s^2 + 2\zeta s + 1}$$
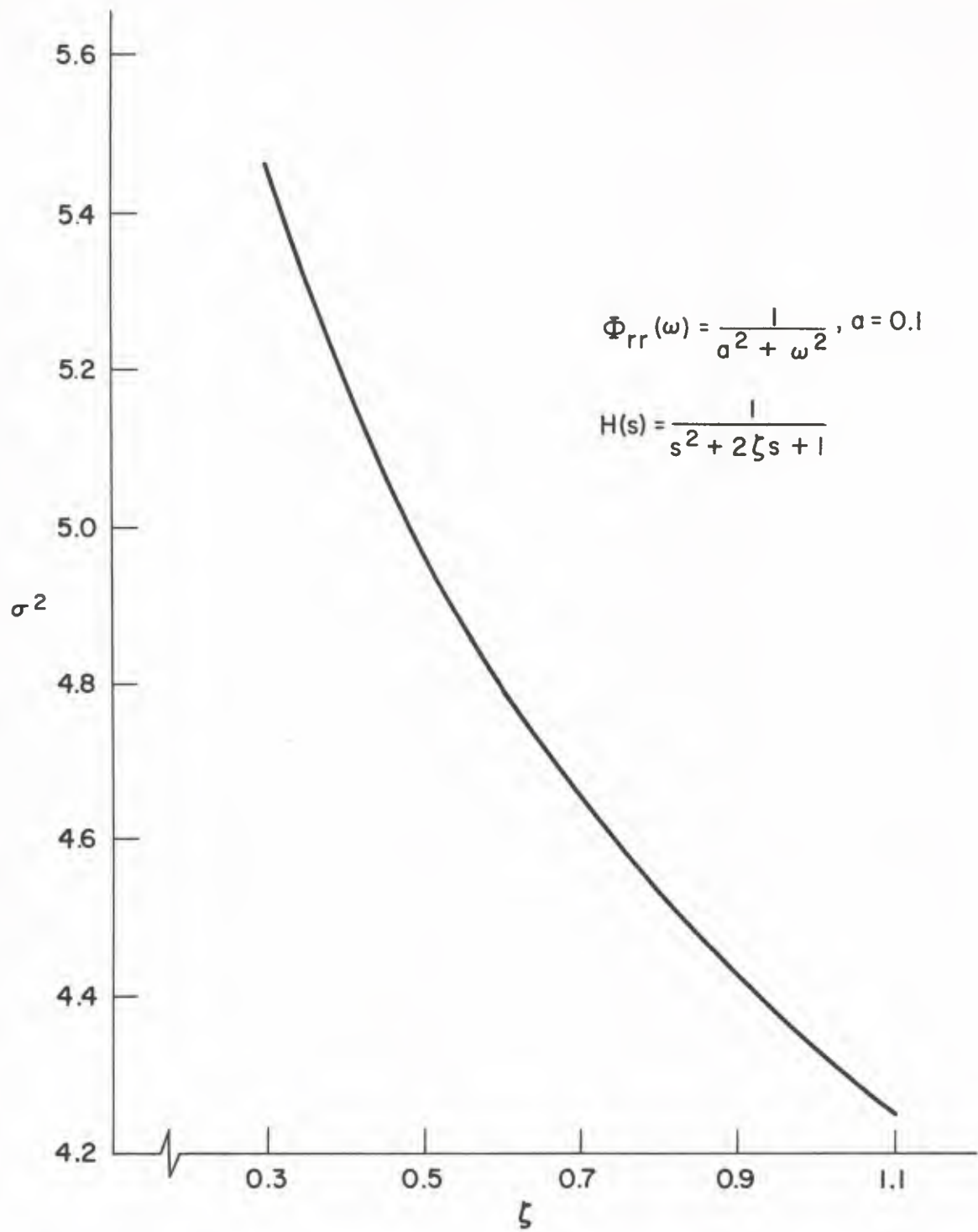
Figure 15. Mean Square Output vs $\zeta$ for a Second-Order Unit Numerator System

48
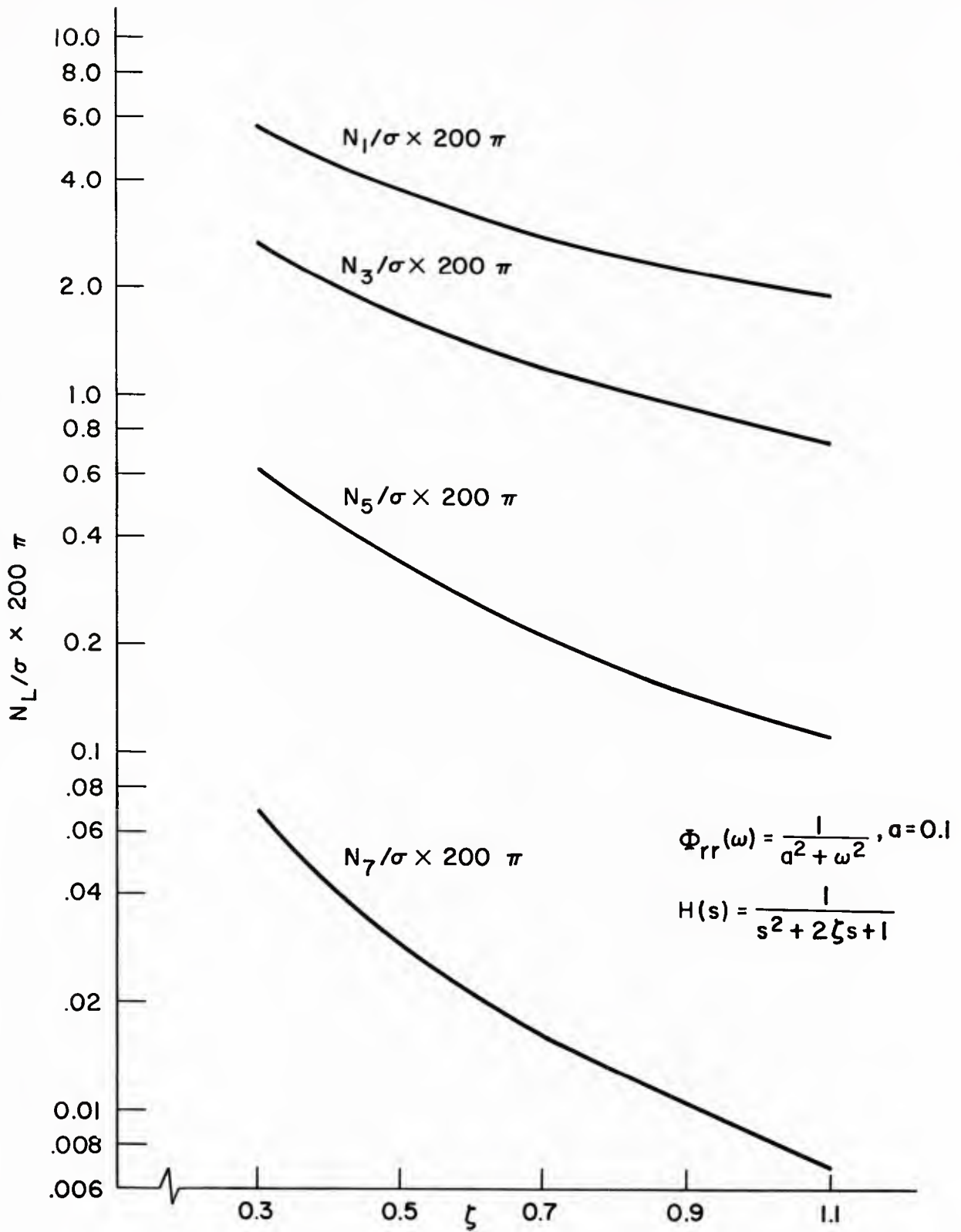
Figure 16. Ratio of Output Exceedances to $\overline{\epsilon^2}$ for Second-Order Unit Numerator System

49

with statistically independent signal and noise input:
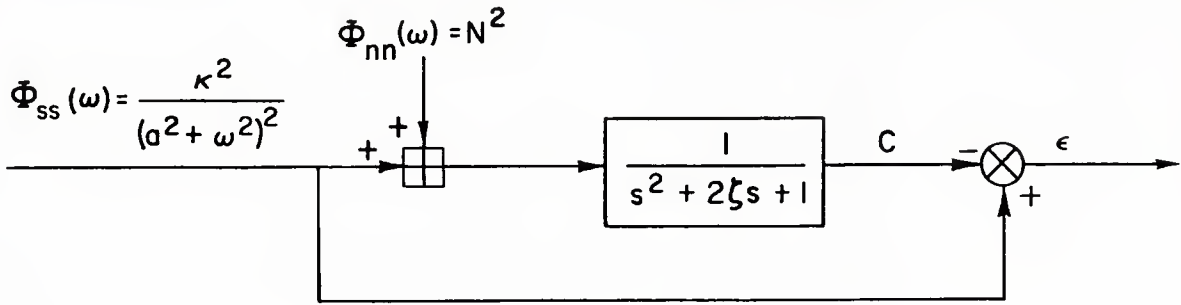


Figure 17. System for Investigation of Error Exceedances

The error is the difference between the signal and the output. The average number of exceedances per second of some specified level of the error, L, is

$$N_{\epsilon_L} = \frac{1}{2} e^{-L^2/2\sigma_\epsilon^2} N_{\epsilon_0} \qquad (53)$$

where $N_{\epsilon_0}$ = the average number of zero crossings per second of the error

$$N_{\epsilon_0} = \frac{1}{\pi} \left[ \frac{\int_0^\infty \omega^2 \Phi_{\epsilon\epsilon}(\omega) d\omega}{\int_0^\infty \Phi_{\epsilon\epsilon}(\omega) d\omega} \right]^{1/2}$$

$$= \frac{1}{\pi} \left[ \frac{\frac{1}{2\pi j} \int_{-j\infty}^{j\infty} -s^2 \Phi_{\epsilon\epsilon}(s) ds}{\frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \Phi_{\epsilon\epsilon}(s) ds} \right]^{1/2}_{s=j\omega}$$

Using the formula for the error power spectrum, $\Phi_{\epsilon\epsilon}(s)$, given in Ref. 50, p. 239,

$$\Phi_{\epsilon\epsilon}(s) = \left[ 1 - H(s) \right] \left[ 1 - H(-s) \right] \Phi_{rr}(s) + H(s)H(-s)\Phi_{nn}(s)$$

50

and substituting the expressions given in Fig. 17 for H, $\Phi_{rr}$, and $\Phi_{nn}$ yields

$$\Phi_{\epsilon\epsilon}(s) = \left[\frac{-s(-s + 2\zeta)}{(s^2 - 2\zeta s + 1)}\right]\left[\frac{s(s + 2\zeta)}{(s^2 + 2\zeta s + 1)}\right]\frac{\kappa^2}{(-s + a)^2(s + a)^2}$$

$$+ \frac{N^2}{(s^2 - 2\zeta s + 1)(s^2 + 2\zeta s + 1)} \tag{54}$$

Note that if the denominator of the input power spectrum had been only of second order (as in the output exceedance example), the numerator of Eq 53 for $N_{\epsilon_0}$ would have failed to converge since $\omega^2\Phi_{\epsilon\epsilon}(\omega)$ would not tend to zero for large $\omega$. This disadvantage is more likely to occur for the error exceedance formulas than for the output exceedance formulas because the order of the numerator of $\Phi_{\epsilon\epsilon}(\omega)$ in Eq 53 is higher than the order of the numerator of the corresponding $\Phi_{cc}(\omega)$. This lack of convergence does not appear to reflect any physical peculiarity of the actual system response, and it should be possible to recast the formulas to avoid this difficulty. This problem may merit further investigation.

Evaluating $N_{\epsilon_0}$ using

$$\sigma_\epsilon^2 \equiv \overline{\epsilon^2} = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty}\Phi_{\epsilon\epsilon}(s)ds$$

yields

$$\overline{\epsilon^2} = \frac{\kappa^2(a + a^2\zeta + 4a\zeta^2 + 4\zeta^3)}{(8a^3 + 4a + 4a^5)\zeta + (16a^4 + 16a^2)\zeta^2 + 16a^3\zeta^3} + \frac{N^2}{4\zeta} \tag{55}$$

and

$$N_{\epsilon_0} = \frac{1}{\pi}\left[\frac{Z}{\overline{\epsilon^2}}\right]^{1/2} \tag{56}$$

$$\text{where } Z = \frac{\kappa^2 2a^2\left[a + a^2(a^2 + 4)\zeta + 4a(a^2 + 1)\zeta^2 + 4a^2\zeta^3\right]}{(8a^3 + 4a + 4a^5)\zeta + (16a^4 + 16a^2)\zeta^2 + 16a^3\zeta^3} + \frac{N^2}{4\zeta}$$
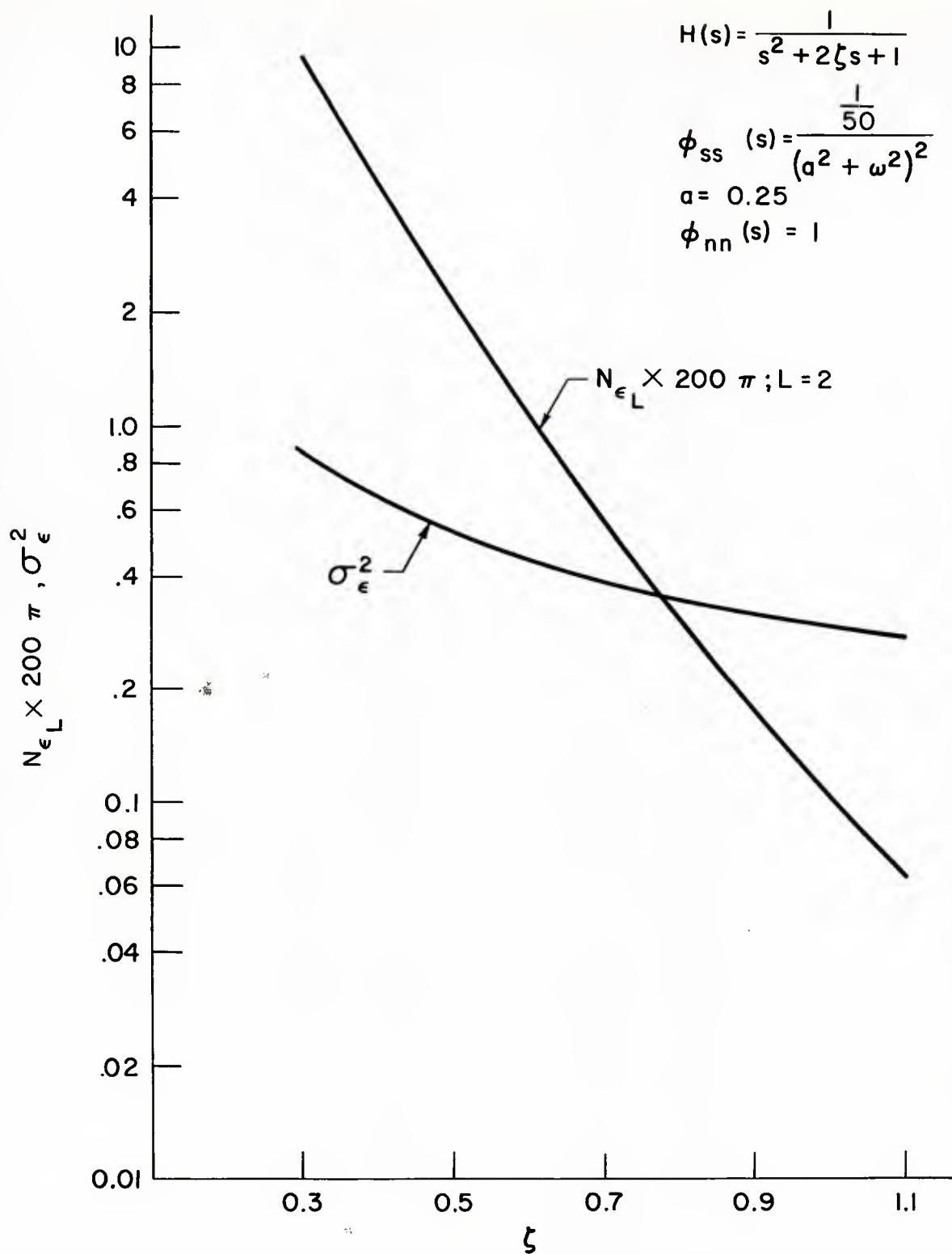
Figure 18.   Error Exceedances and Mean Square Error vs $\zeta$ for a Second-Order Unit Numerator System with 5:1 Low-Frequency Signal-to-Noise Ratio

52
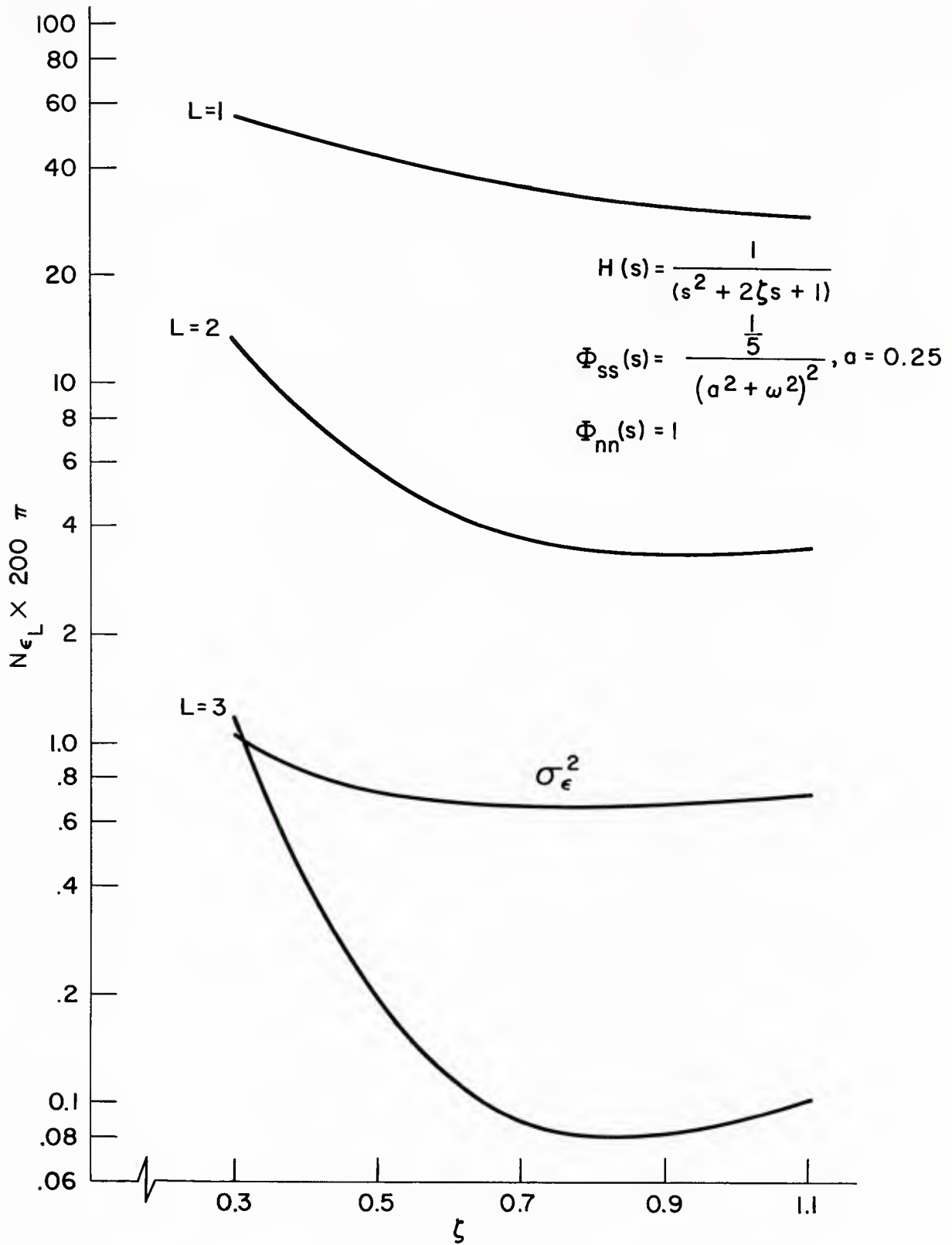
Figure 19. Error Exceedances and Mean Square Error vs $\zeta$ for a Second-Order Unit Numerator System with 50:1 Low-Frequency Signal-to-Noise Ratio

Figure 18 illustrates a typical variation of $N_{\epsilon_L}$ and $\overline{\epsilon^2}$ with $\zeta$ for $L = 2$, $a = 0.25$, $N = 1$, and $K^2 = 1/50$. For this situation the low-frequency ratio of the signal to noise spectra is approximately 5:1. With this signal to noise ratio, neither $\overline{\epsilon^2}$ nor $N_{\epsilon_2}$ yield a satisfactory criterion.

Figure 19 illustrates the effect of changing $K^2$ to $1/5$, thus making the low-frequency signal to noise spectra ratio 50:1. There is a marked change in the behavior of the $N_{\epsilon_L}$ graphs which now display highly selective minima at $\zeta = 0.8$ to 0.9 for large values of $L$.

This change in behavior makes it difficult to form a complete assessment of error exceedance criteria without a much larger number of examples.

## B. PROBABILITY CRITERIA

Ragazzini and Zadeh (Ref. 18) have proposed the probability criterion

$$\text{maximum } p \;=\; \text{maximum probability } |\epsilon| < L$$

for handling situations where it is desirable that the magnitude of the error be less than a certain critical value L; i.e., all errors larger than L are equally bad, while those smaller are equally acceptable. This may well be the case in some problems of ballistics and many flight control applications. It is apparent that for such cases the mean square error does not correctly reflect the requirements of the system. Ragazzini and Zadeh's development of the probability criterion is summarized below; it will be shown that in nontrivial cases, the criterion is difficult to apply.

The problem is set up by Ragazzini and Zadeh in the following way:
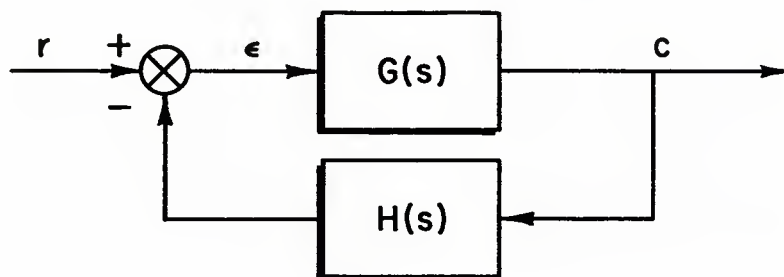


Figure 20. Block Diagram for Probability Criterion Optimization

The error arises from a transient desired input, and from random disturbances arising both inside and outside the system; i.e.,

$$\epsilon = \epsilon_s + \epsilon_n \tag{57}$$

where $\epsilon_s$ is the systematic error due to transient input

$\epsilon_n$ is stationary random noise (normally distributed)

Note that this error is the conventional "actuating error," not the difference between input and desired output

The structure and components of the system are specified completely except for a number of adjustable parameters, $\alpha_1$, $\alpha_2$, $\cdots \alpha_n$. The objective is to assign values to $\alpha_i$ which maximize the probability that at a prescribed instant, $t = t_0$, the magnitude of $\epsilon$ will be less than a given tolerance, L. With a normally distributed noise variable, the probability, p, may be written

$$p = \text{probability} \left[ |\epsilon(t)|_{t=t_0} < L \right] = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{\epsilon_{s0}-L}^{\epsilon_{s0}+L} e^{-u^2/2\sigma^2} du \tag{58}$$

where $\epsilon_{s0}$ is the value of $\epsilon_s$ at $t = t_0$

$\sigma^2$ is the variance of the noise $= E(\epsilon_n^2)$

L is the critical value of $|\epsilon|$

If the sequence of events under consideration is repeated many times (with the same input, the same initial conditions, and the same free parameters), the value of $\epsilon_{s0}$ will be constant. With this understanding, the ensemble average mean square error takes the form

$$E\left[\epsilon^2(t_0)\right] = (\epsilon_{s0})^2 + \sigma^2 \tag{59}$$

In Ref. 18, p is compared with $E\left[\epsilon^2(t_0)\right]$ to examine whether maximizing p also minimizes $E\left[\epsilon^2(t_0)\right]$. From Eq 59, for minimum $E\left[\epsilon^2(t_0)\right]$, it is necessary that

$$\frac{\partial E\left[\epsilon^2(t_0)\right]}{\partial \alpha_i} = 0, \quad i = 1, 2, \cdots n_1 \tag{60}$$

55

Substituting Eq 59 into Eq 60 yields

$$\epsilon_{s0} \frac{\partial \epsilon_{s0}}{\partial \alpha_i} + \sigma \frac{\partial \sigma}{\partial \alpha_i} = 0 \tag{61}$$

Similarly, for p to be a maximum, it is necessary that

$$\frac{\partial p}{\partial \alpha_i} = \frac{\partial p}{\partial \epsilon_{s0}} \left( \frac{\partial \epsilon_{s0}}{\partial \alpha_i} \right) + \frac{\partial p}{\partial \sigma} \left( \frac{\partial \sigma}{\partial \alpha_i} \right) = 0 \tag{62}$$

where i = 1, 2, $\cdots$ n

Choosing the coefficients $\partial p/\partial \epsilon_{s0}$ and $\partial p/\partial \sigma$ in Eq 62 to exactly match those in Eq 61 demands that the ratios $\epsilon_{s0}/(\partial p/\partial \epsilon_{s0})$ and $\sigma/(\partial p/\partial \sigma)$ should be equal. Ragazzini and Zadeh show that this requirement can be satisfied approximately over a wide range of values of $\epsilon_{s0}$, L, and $\sigma$, when

$$|\epsilon_{s0}| \ll \sigma$$

and

$$|L\epsilon_{s0}| \ll \sigma^2$$

or

$$|\epsilon_{s0}| \gg \sigma$$

and

$$|\epsilon_{s0}| \doteq L$$

It is concluded in Ref. 18 that under these conditions, maximizing p is the same as minimizing $E\left[\epsilon^2(t_0)\right]$. However, in all practical situations, it will be found impossible to both satisfy these conditions and achieve a nontrivial result. Either $\epsilon_{s0}$ changes with changes in the $\alpha_i$'s, or it does not. If it does not, all one can do to maximize p is to minimize $\sigma^2$. If $\epsilon_{s0}$ changes with the $\alpha_i$'s, there is no guarantee that the inequality conditions above will always hold as the $\alpha_i$'s are varied. Because this is the most likely case, trial and error methods must be resorted to in order to use the maximum p criterion. Thus, this criterion fails to meet the requirement of ease of application.

C. A NOTE ON THE MINIMUM MEAN SQUARE ERROR CRITERION
   WITH MIXED DETERMINISTIC AND RANDOM INPUT COMPONENTS

The previous section described how probability criteria could be applied to a system subjected to a deterministic command corrupted by random noise. Zadeh and Ragazzini (Ref. 12) have also considered the case where the total signal has three components:

1. $n(t)$, stationary random noise with zero mean

2. $m(t)$, a stationary random signal with zero mean

3. $g(t)$, a deterministic signal representable as a polynomial in $t$ with a finite number of terms.

For this case the minimum mean square error criterion is used with a finite observation time, T; i.e., the task is to find the system that will minimize the ensemble average (over a large number of trials) of the mean square error, the mean of each trial being taken over the observation time, T.

An account of Zadeh and Ragazzini's procedure occupies the whole of Chapter 8 of Ref. 7 ; the procedure is lengthy to explain, complicated, and demands considerable computational labor for problems of practical interest. Indeed, both Ref. 7 and 12 give only simplified examples for which either the random or the deterministic component of the signal is absent. From the viewpoint of the present study —which is directed toward "pencil and paper" methods of optimization — the method fails to meet the requirement for ease of application. In view of this, and because it is difficult to present a detailed account of Zadeh and Ragazzini's procedure that is appreciably shorter than that given in Ref. 7, only an outline of the method is given here.

The presence of the deterministic component in the signal leads to the establishment of certain conditions upon the time moments of the optimum impulsive response. The minimization of the mean square error is then effected by combining these conditions using Lagrange multipliers in a fashion similar to the minimization of performance criteria of systems subject to constraints due to nonlinearities, etc. (e.g., Ref. 6).

Assuming that signal and noise data are available only in the interval 0 to T, the system response is given by

$$c(t) = \int_0^T \left\{ g(t - \tau) + m(t - \tau) + n(t - \tau) \right\} h(\tau)d\tau \qquad (63)$$

where $h(\tau)$ is the actual system weighting function

Since $g(\tau)$, by definition, is expressible as a polynomial in $\tau$ with a finite number of terms, it can be expanded in a Taylor series as

$$g(t - \tau) = g(t) - \tau g'(t) + \frac{\tau^2}{2!} g''(t) + \cdots + (-1)^r \frac{\tau^r}{r!} g^{(r)}(t) \qquad (64)$$

where $r$ is the order of the polynomial describing
the deterministic signal

Substituting Eq 64 into 63 yields

$$c(t) = \mu_0 g(t) - \mu_1 g'(t) + \cdots + (-1)^r \frac{\mu_r}{r!} g^{(r)}(t) \qquad (65)$$

$$+ \int_0^T m(t - \tau)h(\tau)d\tau + \int_0^T n(t - \tau)h(\tau)d\tau$$

where $\quad \mu_\nu = \int_0^T \tau^\nu h(\tau)d\tau, \quad \nu = 0, 1, 2, \cdots r$

Since $m(t)$ and $n(t)$ are stationary with zero means, the ensemble average output at time $(t)$ is only the nonrandom component, $g(t)$.

$$\widetilde{c(t)} = \int_0^T g(t - \tau)h(\tau)d\tau \qquad (66)$$

$$= \mu_0 g(t) - \mu_1 g'(t) + \cdots + (-1)^r \frac{\mu_r}{r!} g^{(r)}(t)$$

If the desired response is physically realizable, then it must be capable of

being related to the total signal by an equation of the form

$$c_d(t) = \int_{-\infty}^{\infty} \left[ g(t - \tau) + m(t - \tau) \right] h_i(\tau) d\tau \qquad (67)$$

where $h_i$ = an "ideal" weighting function
defined by Eq 67

As previously, the stationary random component vanishes when an ensemble average is taken

$$\widetilde{c_d(t)} = \int_{-\infty}^{\infty} g(t - \tau) h_i(\tau) d\tau \qquad (68)$$

For the error to have zero mean,

$$\widetilde{c_d(t)} - \widetilde{c(t)} = 0 \qquad (69)$$

Substituting from Eq 66 and 68 into Eq 69,

$$\int_{-\infty}^{\infty} g(t - \tau) h_i(\tau) d\tau \equiv \mu_0 g(t) - \mu_1 g'(t) + \cdots (-1)^r \frac{\mu r}{r!} g'(t) \qquad (70)$$

Equation 70 determines the values of the first $r + 1$ moments, $\mu_0$, $\mu_1$, $\cdots$ $\mu_r$, of the optimum impulsive response, $h_i(\tau)$. In Ref. 7 it is shown that the ensemble average of the time-averaged square error (i.e., $1/T \int_0^{\infty} \widetilde{\epsilon^2} dt$) can be expressed as

$$\widetilde{\epsilon^2} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_{mm}(\tau - \theta) h_i(\tau) h_i(\theta) d\tau d\theta - 2\int_{-\infty}^{\infty}\int_0^T R_{mm}(\tau - \theta) h_i(\tau) h(\theta) d\tau d\theta$$

$$+ \int_0^T\int_0^T \left[ R_{mm}(\tau - \theta) + R_{nn}(\tau - \theta) \right] h(\tau) h(\theta) d\tau d\theta \qquad (71)$$

where $R_{mm}$ and $R_{nn}$ are the autocorrelation functions of $m(t)$ and $n(t)$
|There is assumed to be zero crosscorrelation between $m(t)$ and $n(t)$|

$\widetilde{\epsilon^2}$ is minimized when $h(t)$ satisfies the following integral equation:

$$\int_0^T \left[ R_{mm}(t - \tau) + R_{mn}(t - \tau) \right] h(\tau) d\tau = \gamma_0 + \gamma_1 t + \cdots + \gamma_r t^r + \int_{-\infty}^{\infty} R_m(t - \tau) h_i(\tau) d\tau$$

$$\text{where } 0 \leq t \leq T \tag{72}$$

$$\gamma_0, \gamma_1, \cdots \gamma_r \text{ are Lagrange}$$
$$\text{multipliers}$$

The procedure for determining $\gamma_0$, $\gamma_1$, $\cdots$ $\gamma_r$ and for solving Eq 71 is described fully in Ref. 7.

The above discussion, and that of the previous section, has been concerned with situations where the command is partially or completely deterministic. This implies well-defined initial conditions (at least within certain limits) and often a finite time of observation. These conditions apply to such problems as missile interception, orbital rendezvous, automatic landing control, etc. For these problems, terminal errors are much more important than errors occurring earlier. Therefore, it seems more logical to employ time-weighted criteria (instead of minimum $\overline{\epsilon^2}$) so that the analytical optimization procedure truly reflects how the importance of the flight control system function varies with time. Such time-weighted criteria are discussed in the chapter that now follows.

60

# CHAPTER III

## TIME-WEIGHTED CRITERIA

Time-weighted criteria are very suitable for normalized linear constant-coefficient systems with deterministic inputs. For example, criteria such as

$$\text{minimum ITAE} \quad = \quad \min \int_0^\infty t\,|\epsilon|\,dt$$

$$\text{minimum IT}^3\text{E}^2 \quad = \quad \min \int_0^\infty t^3[\epsilon]^2\,dt$$

have been shown (Ref. 1) to yield satisfactory systems for step inputs, and are highly selective. It is logical to inquire whether corresponding criteria could be evolved for random inputs. This chapter investigates this problem by examining the three published criteria listed below.

1. Murphy and Bold's Criterion

$$\text{minimum} \quad \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{+T} W(t)\left[\epsilon(t)\right]^2 dt$$

where $W(t)$ is a time-weighting function statistically independent of $f[\epsilon(t)]$ and may be deterministic

2. Zaborszky and Diesel's Probabilistic Square Error Criterion

$$\text{minimum} \quad \int_0^\infty \widetilde{p(t)\epsilon^2(t)}\,dt$$

where $\widetilde{\epsilon^2(t)}$ is the ensemble average of system error squared taken over a number of trials, and $p(t)$ can be interpreted as a time-weighting function indicating the relative importance of errors occurring at different times. It is usually permissible to assume the trials to be concurrent; the averaging sign then need not be extended over $p(t)$.

61

3. Glover's Criterion

$$\text{minimum} \quad \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} \frac{\epsilon^2(t)}{\left[c_d(t)\right]^2 + \delta^2} \, dt$$

where $c_d(t)$ is the desired output, and $\delta$ is a
constant which dictates the lowest absolute
accuracy of interest

The principal conclusions of this chapter are that

1.  Murphy and Bold's criterion yields the same optimum system for
    all choices of $W(t)$, if the input environment is stationary.
    This is true even when the criterion is generalized by replac-
    ing $\overline{\epsilon^2}$ with any other $f(\epsilon)$

2.  the probabilistic square error criterion is suitable for
    nonstationary situations (such as repeated missile trials)
    when a meaningful selection of $p(t)$ can be made

3.  Glover's criterion, when applied to free-order linear systems
    with uncorrelated Gaussian input and noise, yields optima
    closely related to the Wiener system.  When the system order
    is constrained (i.e., Phillips-type optimization), the criterion
    tends to be unselective.

These conclusions are summarized in Fig. 21.

A.  MURPHY AND BOLD'S CRITERION

In Ref. 20, Murphy and Bold propose the MWSE (mean weighted square error)
criterion

$$\text{minimum MWSE} \quad = \quad \text{minimum} \quad \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} W(t)\epsilon^2(t)dt \tag{73}$$

$W(t)$ is statistically independent of $\epsilon(t)$, and may be deterministic.  This crite-
rion is a special case of the more general criterion

$$\text{minimum MWFE} \quad = \quad \text{minimum} \quad \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} W(t)f\left[\epsilon(t)\right]dt \tag{74}$$

62

GENERAL FORMS

CRITERION

Minimum $\overline{W(t)f[\epsilon(t)]}$

Minimum $\int_0^\infty p(t)f[\epsilon(t)]dt$

SPECIAL
CONDITIONS

W(t) statistically
independent of $\epsilon(t)$.
[W(t) may be random
or deterministic.]

W(t) statistically dependent
on $\epsilon(t)$, e.g., Glover's
amplitude-weighted measure,

$W(t) = \dfrac{1}{[c_d(t)]^2 + \delta^2}$

INPUT
ENVIRONMENT

Stationary input

Nonstationary input

Stationary input

CONCLUSIONS

As shown on p. 65
MWFE = $\overline{W(t)}$ x $\overline{f[\epsilon(t)]}$.
Hence, minimum $\overline{f[\epsilon(t)]}$
system is optimum. If
$f[\epsilon(t)] = [\epsilon(t)]^2$, then
the optimum system is
Wiener or Phillips-type.

Valid criterion if meaning-
ful p(t) can be specified.
Optimization using $f[\epsilon(t)] = \epsilon^2$
has been extensively studied
and is (in principle) simple
to perform.

With uncorrelated Gaussian
signal and noise, the opti-
mum system is a constant
times the Wiener system. No
similar simple relationship
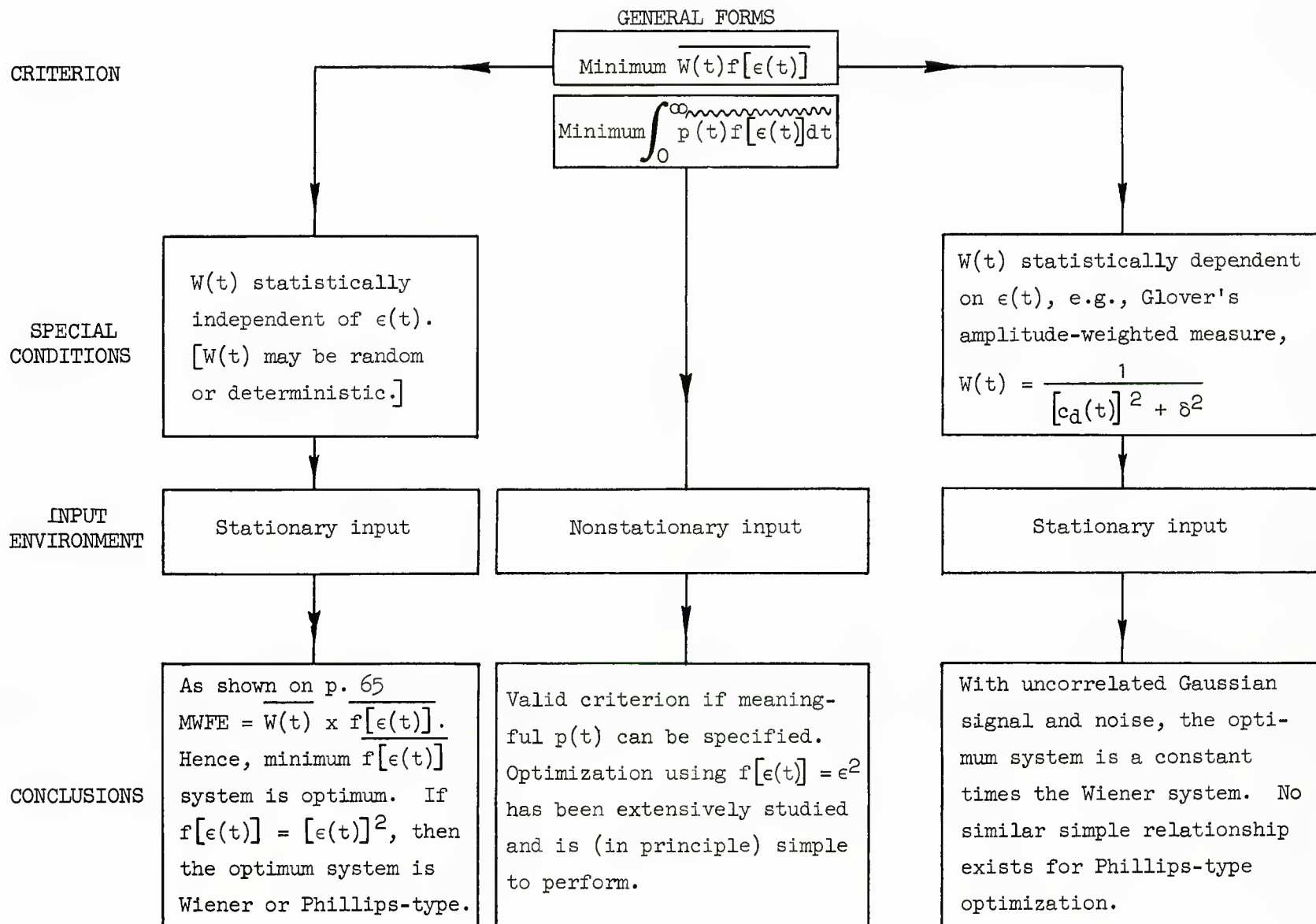exists for Phillips-type
optimization.

Figure 21. Summary of Time-Weighted Performance Measures

Although it was not investigated by Murphy and Bold, many of the results to be proved below also apply to this form.

Murphy and Bold derive the expression for the optimum MWSE system by a procedure analogous to Wiener's. A sufficient condition for the MWSE to be a minimum is also derived, i.e.,

$$W(t) > 0 \quad \text{for} \quad -\infty < t < \infty$$

which one would intuitively expect, because this keeps the integrand in Eq 74 positive. Murphy and Bold apparently failed to realize that minimizing the MWSE is equivalent to minimizing $\overline{\epsilon^2}$ (which will be proved below). This point also seems to have been overlooked by Zaborszky and Diesel (Ref. 49) in their comments on Ref. 20, and in the accompanying reply of Murphy and Bold. The equivalence of minimizing MWSE and $\overline{\epsilon^2}$ is proved by demonstrating that

$$\text{MWSE} = \left(\overline{\epsilon^2}\right)\left(\overline{W}\right) \tag{75}$$

Proof

By definition, $\qquad\qquad \text{MWSE} = \overline{W(t)\epsilon^2(t)} \tag{76}$

where the bar denotes time averaging

Consider $W(t)$ to be deterministic with a finite time-average value and $\epsilon(t)$ to have zero mean. Taking the ensemble average of a number of repeated trials yields (using the same system for each trial)

$$E\big[\text{MWSE}\big] = E\left[\overline{W(t)\epsilon^2(t)}\right] \tag{77}$$

As shown on p. 65 of Ref. 50, it is permissible to interchange the order of time and ensemble averaging for this expression. A series of stationary inputs, each having the same spectrum, when applied to identical systems yields the same $\overline{\epsilon^2}$ for each system. For the MWSE criterion to be of practical use, it is likewise necessary that when a series of stationary inputs, each having the same spectrum, is applied to identical systems the MWSE values for each system must be equal. Thus $E|\text{MWSE}| = \text{MWSE}$ and Eq 77 becomes

$$\text{MWSE} = \overline{E\big[W(t)\epsilon^2(t)\big]} \tag{78}$$

Since $W(t)$ is constant when averaging across the ensemble at any particular value of $t$, then

$$E\left[W(t)\epsilon^2(t)\right] \;=\; W(t)E\left[\epsilon^2(t)\right] \;=\; W(t)\,\sigma_\epsilon^2 \tag{79}$$

Because (for a stationary input environment) $\sigma_\epsilon^2$ is a constant, substituting Eq 79 into Eq 78 yields

$$MWSE \;=\; \overline{W(t)\sigma_\epsilon^2} \;=\; \sigma_\epsilon^2(\overline{W}) \tag{80}$$

Thus, minimizing MWSE for any specified deterministic $W(t)$ is simply equivalent to minimizing $\sigma_\epsilon^2$.

For the case where $W(t)$ is a random signal statistically independent of $\epsilon(t)$, the proof proceeds as indicated above, except that Eq 78 leads directly to

$$MWSE \;=\; \overline{E(W\epsilon^2)} \;=\; \overline{E(W)E(\epsilon^2)} \;=\; \overline{(\overline{W})(\overline{\epsilon^2})} \;\equiv\; (\overline{W})(\overline{\epsilon^2}) \tag{81}$$

because the ensemble average of the product of two statistically independent variables is merely the product of their individual ensemble averages.

It can also be shown that minimizing the more general MWFE criterion is equivalent to minimizing the appropriate $\overline{f(\epsilon)}$, because (by a similar proof to that given above)

$$MWFE \;=\; \left[\overline{f(\epsilon)}\right](\overline{W}) \tag{82}$$

It was shown in Chapter II that, for many input distributions, minimization of the very broad class of $\overline{f(\epsilon)}$ measures that are even and nondecreasing with $|\epsilon|$ is equivalent to minimizing $\overline{\epsilon^2}$. Combining this result with Eq 82 leads to the observation that in very many cases the MWFE criterion simply yields the minimum $\overline{\epsilon^2}$ system.

B.  THE "END-SIGMA" AND "PROBABILISTIC SQUARE ERROR" CRITERIA

Zaborszky and Diesel first proposed the end-sigma performance measure in
Ref. 22, and developed it further in Ref. 23, 24, 25, 26, and 27.  This measure
is defined as

$$S = \int_0^\infty F\overbrace{[\epsilon(t), \; t, \; \nu_1, \; \cdots \; \nu_r}]p(t)dt \tag{83}$$

where the wavy line indicates averaging over
an ensemble of inputs of different types

The integrand contains a function of error, time, and various system parameters
$\nu_1$, $\nu_2$, $\cdots$ $\nu_r$, and $p(t)$, the probability that the output will be used at
time t.  It is claimed (Ref. 22) that "this measure unites in a single concept
the transient and steady states of the system operation as well as the largely
neglected intermediate states.  None of these operating states is discriminated
for or against, and because of $p(t)$, each gets the weight due it because of the
relative frequency of its presence at such times as the output is utilized."  The
generality of the measure makes it impossible to dispute this claim, and credit is
certainly due to Zaborszky and Diesel for recognizing the need for a criterion
which is compatible with both deterministic and random inputs.  However, to use
the criterion, it must first be put into a concrete form.  The particular concrete
forms chosen by Zaborszky and Diesel will now be discussed, and validity, selec-
tivity, and ease of application as performance criteria will be assessed.  In each
of Ref. 22, 23, 24, 25, 26, and 27 the measure considered was

$$S = \int_0^\infty \overbrace{[\epsilon(t)]}^2 p(t)dt \tag{84}$$

where $p(t)$ is independent of $\epsilon(t)$

which Zaborszky and Diesel call the "probabilistic square error."  This form is
meaningful only for nonstationary situations since for stationary problems the
probabilistic square error must be expressed as

66

$$\varsigma = \int_0^\infty p(t) \widetilde{\left[\epsilon(t)\right]^2} dt = \lim_{T \to \infty} \int_0^T p_1(t) \widetilde{\left[\epsilon(t)\right]^2} dt \qquad (85)$$

where $p(t) = \lim_{T \to \infty} \frac{1}{T} \left[ u(t) - u(t - T) \right] p_1(t)$,

and $p_1(t)$ is a general probability distribution function

If $p_1(t)$ is independent of $\epsilon(t)$, where $\epsilon(t)$ is a random function of time, minimizing $\varsigma$ is equivalent to minimizing the mean square error (as has been shown in the discussion of Murphy and Bold's criterion in Section A of this chapter.

Thus the probabilistic square error criterion might be appropriate for non-stationary situations such as automatic landing systems or missile flight control systems in which the control system is required to operate for limited durations, and in which the ensemble average performance over a number of trials constitutes the basis for system assessment.

Evaluation of the Probabilistic Square Error in the s-Domain

Two procedures for evaluating $\varsigma$ in the s-domain will now be given. First the procedure suggested by Zaborszky and Diesel in Ref. 22 will be summarized, and then an alternative procedure which is often more convenient for simple systems will be described.

It is shown in Ref. 22 that when each member of the ensemble is the same, $\widetilde{\left[\epsilon(t)\right]^2} = \left[\epsilon(t)\right]^2$ and the probabilistic square error can be rephrased as

$$\varsigma = \lim_{t \to 0} \int_0^\infty p(\tau - t) \left[\epsilon(\tau)\right]^2 d\tau \qquad (86)$$

$$= \lim_{t \to 0} \mathcal{L}^{-1} \left[ P(-s)E(s)^*E(s) \right] \qquad (87)$$

Numerator and denominator polynomials are defined such that

$$P(-s) = \frac{A(s)}{B(s)} \qquad (88)$$

and

$$E(s)^*E(s) \quad = \quad \frac{J(s)}{L(s)} \tag{89}$$

and it is demonstrated that

$$\varsigma = \sum_i \frac{A(p_i)}{B(p_i)} \frac{J(p_i)}{\left[dL/ds\right]_{s=p_i}} \tag{90}$$

where $p_i$ are the zeros of $L(s)$, all of which are assumed to be simple and located in the left half-plane

Alternative formulations of $\varsigma$ are given where these conditions cannot be satisfied (Ref. 22).

Example of Probabilistic Square Error Evaluation

The following example is presented by Zaborszky and Diesel to demonstrate the computation process in general: a system described by the closed-loop transfer function

$$\frac{C(s)}{R(s)} \quad = \quad \frac{s+1}{s^2+s+1} \tag{91}$$

is subjected to an input, $r(t)$, where

$$r(t) \quad = \quad 1 + 2t \tag{92}$$

The probability function of the output is

$$p(t) \quad = \quad \frac{10}{3} (e^{-2t} - e^{-5t}) \tag{93}$$

This example was solved in Ref. 22 using the procedure described above, which is quite general but requires the zeros of $L(s)$ to be known. However, alternative methods are possible for an example of such simplicity. (The procedure of Mishkin and Braun, p. 314 of Ref. 51, is worthy of mention; but this also requires knowledge of the zeros of $L(s)$.)

68

A method will now be demonstrated which has the advantage of obviating the need to find these zeros. Substituting the $p(t)$ of Eq 93 into Eq 84 for $\varsigma$

$$\varsigma = \frac{10}{3} \int_0^\infty e^{-2t} [\epsilon(t)]^2 dt - \frac{10}{3} \int_0^\infty e^{-5t} [\epsilon(t)]^2 dt \tag{94}$$

$$= \frac{10}{3} \left[ \mathcal{L}[\epsilon(t)]^2 \right]_{s=5}^{s=2} \tag{95}$$

Stone (Ref. 52) has given general literal forms for $\mathcal{L}\left[\mathcal{L}^{-1} E(s)\right]^2$. In particular,

$$\mathcal{L}\left[\mathcal{L}^{-1} \frac{B_1 s + B_0}{s^2 + A_1 s + A_0}\right]^2 = \frac{B_1^2 s^2 + (B_1^2 A_1 + 2B_0 B_1) s + (2B_1^2 A_0 + 2B_0^2)}{s^2 + 3A_1 s^2 + (2A_1^2 + 4A_0)s + 4A_0 A_1} \tag{96}$$

So, in the present example,

$$E(s) = \left|\frac{C(s)}{R(s)}\right| R(s) = \frac{s^2}{s^2 + s + 1} \cdot \frac{s + 2}{s^2} = \frac{s + 2}{s^2 + s + 1} \tag{97}$$

Hence the probabilistic square error is

$$\varsigma = \frac{10}{3} \left[ \frac{s^2 + 5s + 10}{s^2 + 3s^2 + 6s + 4} \right]_{s=5}^{s=2} = \frac{10}{3} \frac{24}{36} - \frac{60}{234} = 1.37 \tag{98}$$

which agrees with the result given in Ref. 22. In addition to the fact that there is no need to evaluate the zeros of $L(s)$ using Stone's formulas, this procedure may also be more convenient for minimization, because general expressions for $\partial\varsigma/\partial B_1$, $\partial\varsigma/\partial B_0$, etc., can be derived. However, for higher-order systems Stone's formulas become exceedingly long, and Zaborszky and Diesel's procedure is to be preferred, particularly if a digital computer is available for determining the zeros of $L(s)$.

## Minimization of the Probabilistic Square Error in the Time Domain

In Ref. 22 and 23, Zaborszky and Diesel discuss the evaluation and minimization of probabilistic square error using time-domain operations. There is some correspondence between the given procedure and Wiener optimization, but the methods adopted by Zaborszky and Diesel to solve the equation corresponding to the Wiener-Hopf equations are essentially approximate, since the optimum weighting function is expressed as

$$\sum_{i=0}^{n} k_i t^i$$

where n is predetermined

The accuracy of the solution for the optimum weighting function is thus limited by the largest value of n that can be handled using the computational aids available. However, time-domain synthesis is advantageous in the following respects:

1. nonexponential forms for $p(t)$ can be handled without analytical difficulties. (In the example to be discussed, $p(t) = 1/s$ for $1 < t < 6$, and is zero outside this range.) As noted on p. 313 of Ref. 51, such a form would be analytically inconvenient for the s-domain procedures discussed above

2. the physical significance of each step in the optimization procedure is readily apparent.

For convenience of digital computer programming, in Ref. 22 the input, desired output, probability weighting function, etc., are approximated by sums of Legendre polynomials. It will be demonstrated below that such a representation is not essential, and that the example of Ref. 22 can be worked by time-domain procedures not (explicitly) involving Legendre polynomials. This simplification enables the physical interpretation of the time-domain approach to be kept clearly in sight throughout the entire optimization procedure, thus fully exploiting the advantages of time-domain synthesis. The error can be defined as

$$\epsilon(t) = r(t) - \int_0^\infty h(t_1) r(t - t_1) dt_1 \tag{99}$$

where $h(t_1)$ is the system weighting function
(k is used rather than h to facilitate
reference to Ref. 22)

70

With

$$\varsigma = \int_0^\infty p(t)\,\overbrace{\epsilon^2(t)}\,dt \tag{100}$$

it can be shown (Ref. 23) that

$$= \psi_{c_d c_d}(0,0) - 2\int_0^\infty h(t_1)\psi_{c_d r}(0,-t_1) + \int_0^\infty h(t_3)dt_3\int_0^\infty h(t_1)\psi_{rr}(t_3,\,t_3 - t_1)dt_1 \tag{101}$$

$$\text{where}\quad \psi_{rr}(t,\,\tau) = \int_0^\infty \overbrace{p(t_2 + t)r(t_2)r(t_2 + \tau)}\,dt_2$$

$$\psi_{c_d c_d}(t,\,\tau) = \int_0^\infty \overbrace{p(t_2 + t)c_d(t_2)c_d(t_2 + \tau)}\,dt_2$$

$$\psi_{c_d r}(t,\,\tau) = \int_0^\infty \overbrace{p(t_2 + t)c_d(t_2)r(t_2 + \tau)}\,dt_2$$

The minimization of $\varsigma$ is effected by equating its first variation, $\delta\varsigma$, to zero, which leads to

$$\int_0^\infty h(t_1)\psi_{rr}(t,\,t - t_1)dt_1 = \psi_{c_d r}(0,-t) \tag{102}$$

The resemblance of Eq 102 to the Wiener-Hopf equation should be noted. However, in the Wiener-Hopf equation, the term corresponding to $\psi_{rr}$ is a function of $t - t_1$ only, so the techniques that have been evolved for solving the Wiener-Hopf equation cannot be applied to Eq 102.

Zaborszky and Diesel solve Eq 102 by approximating k, $\psi_{rr}$, and $\psi_{c_d r}$ with sums of orthogonal functions such as Legendre polynomials. This yields n simultaneous linear equations, where n is the order of the highest Legendre polynomial employed. In general, the "optimum" weighting function so obtained is approximate, in that the accuracy of the approximation is limited by the magnitude of n, which in turn is limited by the capacity of the computational aids available.

71

An alternative time-domain procedure for obtaining an $n^{th}$-order polynomial approximation to the optimum weighting function will now be outlined. The procedure is most easily understood by considering an example, and the problem given in Ref. 23 will be used for this purpose.

The example discussed in Ref. 23 consists of the optimization of a system having a weighting function, $h(t)$, such that the output will be a prediction of the input 1 sec from the present time. The input is

$$r(t) = u(t)\left[1 + 0.5t + 2t^2\right] \tag{103}$$

The desired output, $c_d$, is thus $r(t + 1)$, which becomes

$$c_d(t) = u(t)\left[3.5 + 4.5t + 2t^2\right] \tag{104}$$

The probabilistic square error criterion is used with equal weight attached to all outputs occurring between $t = 1$ and $t = 6$ sec, and zero weighting is given to outputs occurring outside this period, i.e., $p(t) = (1/5)\left[u(t - 1) - u(t - 6)\right]$. Hence, the quantity to be minimized becomes

$$s = \int_0^\infty \frac{1}{5}\left[u(t - 1) - u(t - 6)\right]\left[c(t) - c_d(t)\right]^2 dt \tag{105}$$

The actual output, $c(t)$, is given by the convolution

$$c(t) = \int_0^t h(\tau)r(t - \tau)d\tau \tag{106}$$

Following Ref. 23, $h(t)$ will be approximated by the expression

$$h(t) = k_0 + k_1 t + k_2 t^2 \tag{107}$$

Strictly interpreted, this would give an unstable weighting function, but because the probability that the output will be used is zero after $t = 6$ sec, this point is of academic interest only.

Performing the convolution,

$$c(t) = \int_0^t (k_0 + k_1\tau + k_2\tau^2)\left[1 + 0.5(t - \tau) + 2(t - \tau)^2\right] d\tau$$

$$= k_0 t + \left(\frac{k_0}{4} + \frac{k_1}{2}\right)t^2 + \left(\frac{2k_0}{3} + \frac{k_1}{12} + \frac{k_2}{3}\right)t^3 + \left(\frac{k_1}{6} + \frac{k_2}{24}\right)t^4 + \left(\frac{k_2}{15}\right)t^5 \quad (108)$$

A necessary condition for $\varsigma$ to be a minimum is that

$$\frac{\partial}{\partial k_n} \int_1^6 \left\{-3.5 + \left(k_0 - 4.5\right)t + \left(\frac{k_0}{4} + \frac{k_1}{2} - 2\right)t^2\right.$$

$$\left. + \left(\frac{2}{3}k_0 + \frac{k_1}{12} + \frac{k_2}{3}\right)t^3 + \left(\frac{k_1}{6} + \frac{k_2}{24}\right)t^4 + \left(\frac{k_2}{15}\right)t^5\right\}^2 dt = 0 \quad (109)$$

which must be satisfied for $n = 0$ and $1$ and $2$

This is simply evaluated by noting that the differentiation can be brought under the integral sign. For example, with $n = 0$, Eq 109 is equivalent to

$$\int_1^6 2\left\{-3.5 + \left(k_0 - 4.5\right)t + \left(\frac{k_0}{4} + \frac{k_1}{2} - 2\right)t^2 + \left(\frac{2}{3}k_0 + \frac{k_1}{12} + \frac{k_2}{3}\right)t^3\right.$$

$$\left. + \left(\frac{k_1}{6} + \frac{k_2}{24}\right)t^4 + \left(\frac{k_2}{15}\right)t^5\right\}\left\{t + \frac{t^2}{4} + \frac{2}{3}t^3\right\} dt = 0 \quad (110)$$

This gives the solution of $\partial\varsigma/\partial k_0 = 0$ in terms of $k_0$, $k_1$, and $k_2$.

Evaluating this integral and the corresponding equations for

$$\frac{\partial\varsigma}{\partial k_1} = 0, \qquad \frac{\partial\varsigma}{\partial k_2} = 0$$

73

leads to a set of three simultaneous equations for $k_0$, $k_1$, and $k_2$. The corresponding equations derived by Zaborszky and Diesel are, in matrix form,

$$
\begin{bmatrix}
4,392.06 & -4,165.43 & 1,412.35 \\
-15,206.65 & 13,815.24 & -4,543.13 \\
3,538.83 & 3,188.88 & 1,335.96
\end{bmatrix}
\begin{bmatrix}
k_0 \\
k_1 \\
k_2
\end{bmatrix}
=
\begin{bmatrix}
8,829.85 \\
-29,021.61 \\
7,912.02
\end{bmatrix}
\tag{111}
$$

These equations are mildly ill-conditioned; this becomes more apparent when the coefficients of $k_0$ are made unity. The matrix is then

$$
\begin{bmatrix}
1 & -0.964 & +0.3215 \\
1 & -0.908 & +0.2987 \\
1 & -0.901 & +0.3775
\end{bmatrix}
\begin{bmatrix}
k_0 \\
k_1 \\
k_2
\end{bmatrix}
=
\begin{bmatrix}
2.01 \\
1.908 \\
2.236
\end{bmatrix}
\tag{112}
$$

For a weighting function as simple as that considered in Eq 105, the ill-conditioning is unlikely to be troublesome, but it will probably become more pronounced as the order of the approximation to $h(t)$ is increased. However, higher-order weighting function approximations will require digital computers to evaluate the simultaneous equations, and the availability of digital computers will ease the difficulties associated with ill-conditioning.

The time-domain synthesis procedure presented in Ref. 23 is essentially directed towards digital computer evaluation, and differs from the technique described above in that $p(t)$, $r(t)$, $c_d(t)$, and $h(t)$ are represented by sets of orthogonal functions. The resulting expressions are fairly complicated, but can be split into two parts, one of which is invariant with $p(t)$, $r(t)$, and $c_d(t)$, so that this part can be retained throughout a range of problems. The entire programming procedure is tabulated in Ref. 25, and an example is given. An extended description of how a homing aircraft flight control system can be reduced to a form suitable for $\varsigma$-optimization is given in Ref. 27. In Ref. 26, a continuous control system is optimized subject to a constraint of acceleration limiting on the output, $c(t)$. The resulting nonlinearity is avoided by the artifice of constraining the rms

74

output of a linear substitute system (actually the original system with the non-linearity removed) through the use of Lagrange multipliers. In Ref. 24 the probabilistic square error of a sampled-data system is considered, and extensive tabulations are given of the required digital programming procedure.

Assessment of the Probabilistic Square Error Criterion

It is hardly possible to assess such a generalized performance measure as

$$\int_0^\infty \overbrace{p(t) \, \overline{e^2}(t) dt}$$

without making a number of arbitrary choices for $p(t)$. In the present generalized study, no sufficient basis exists for such a choice, but it is at least arguable that in specific applications this choice will be easily made. Accepting this premise (together with the consequence already mentioned that for some $p(t)$, unstable systems may result from the optimization procedure), it appears that the probabilistic square error criterion and its generalized form, the end-sigma criterion, are quite suitable for the optimization of nonstationary systems. The present report has attempted to indicate ways in which the criterion may be handled using desk rather than digital computers. In particular, a suggested procedure for s-domain synthesis of low-order systems has been discussed. For the analysis of flight control problems, it is usually permissible to replace the actual flight control system by a low-order equivalent system (page 7), and the use of this substitution should render Zaborszky and Diesel's criterion suitable for "pencil and paper" studies.

C. GLOVER'S AMPLITUDE-WEIGHTED PERFORMANCE MEASURE

In Ref. 21 Glover reasoned that "the criterion of least mean square error has the disadvantage that large effects are weighted quite heavily even when they occur at a time when the variable under consideration is large." He further reasoned that usually the error expressed as a percentage of the desired output is of more interest than is the absolute error, and therefore proposed a performance measure in which the error time-weighting function contains the amplitude of the desired output. For convenience, the performance measure will be referred to as P.E. because the integrand involves error expressed as a percentage (or

fraction) of the desired output:

$$P.E. = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \frac{\epsilon^2(t)}{\left[c_d(t)\right]^2 + \delta^2} \, dt \tag{113}$$

where  $\epsilon(t)$ is the error

$c_d(t)$ is the desired output

$\delta$ is a constant which dictates the lowest absolute accuracy of interest. It prevents the integrand from becoming infinite whenever $c_d(t) = 0$.

Glover used the criterion of minimum P.E. with the restriction that only linear time-invariant systems would be considered. He showed that with uncorrelated Gaussian signal and noise, the optimum linear system is merely the Wiener system multiplied by a constant. Glover's analysis is outlined below, and an extension is then derived for fixed-order (Phillips-type) optimization. The principal conclusions of this section are:

1.  because the Glover optimum linear system is so simply related to the Wiener system, it suffers from the possible disadvantages of the latter, e.g., poor transient response

2.  for Phillips-type optimization, Glover's performance measure tends to be even less selective than the minimum mean square error criterion.

General Expression for P.E. and the Optimum Linear System

Glover's procedure for the optimum linear system will be outlined briefly below, while simultaneously manipulating P.E. into a form suitable for Phillips-type optimization.

The expression for P.E. (Eq 113) can be written

$$\text{P.E.} \quad = \quad \overline{\left[ \frac{c_d^2 - 2c_d c + c^2}{c_d^2 + \delta^2} \right]} \tag{114}$$

where $\quad c_d \quad = \quad f(t + \eta)$

$$c \quad = \quad \int_0^\infty h(x)r(t - x)\,dx$$

$h(x) \quad =$ the impulsive response of a general linear physically realizable system

Now let $W(t) = \dfrac{1}{\left[ f(t + \eta)\right]^2 + \delta^2}$ be the weighting function on $\epsilon^2$ (Eq 113). Substituting into Eq 114, the expressions for $c_d$, $c$, and $W$ yield

$$\text{P.E.} \quad = \quad \overline{c_d^2 W} - 2\int_0^\infty h(x)\,\overline{r(t - x)f(t + \eta)W(t)}\,dx$$

$$+ \int_0^\infty \int_0^\infty h(x)h(y)\,\overline{r(t - x)r(t - y)W(t)}\,dx\,dy \tag{115}$$

Glover's definition of a "weighted correlation function,"

$$R_{Wuv}(\tau_1, \tau_2) \quad = \quad \overline{u(t + \tau_1)v(t + \tau_2)W(t)} \tag{116}$$

allows P.E. (Eq 115) to be expressed compactly as

$$\text{P.E.} \quad = \quad R_{Wff}(0, 0) - 2\int_0^\infty h(x)R_{Wrf}\left[-(x + \eta), \, 0\right]dx$$

$$+ \int_0^\infty \int_0^\infty h(x)h(y)R_{Wrr}\left[-(x + \eta), \, -(y + \eta)\right]dx\,dy \tag{117}$$

77

Variational calculus methods, as in the Wiener derivation, are used by Glover to establish that for P.E. to be a minimum, it is necessary that $h(x)$ satisfy the integral equation

$$R_{Wrf}[-(x + \eta),\ 0] = \int_0^\infty h(y)R_{Wrr}[-(x + \eta),\ -(y + \eta)]dy \qquad (118)$$

Reference 21 indicates that the solution of the integral equation would be immensely simplified if the weighted correlation function could be expressed in terms of the ordinary correlation functions. This will now be accomplished using the ergodic theorem to express $R_{Wuv}(\tau_1,\ \tau_2)$ in terms of ensemble averages.

Determination of the Optimum Weighting Function
for Gaussian Input and Noise

Glover expresses the weighted correlation function as ensemble averages by the relation

$$R_{Wuv}(\tau_1,\ \tau_2) = \iiint_{-\infty}^\infty \frac{uv}{c_d^2 + \delta^2}\ p(u,\ v,\ c_d)dudvdc_d \qquad (119)$$

> where $p(u,\ v,\ c_d)$ is the joint probability density function, and $u$ and $v$ are the general variables used in Eq 116

Substituting this form of weighted crosscorrelation function into Eq 118 leads to solutions for the optimum weighting function when the distributions involved can be specified analytically. For Gaussian input environments, there are three cases of interest:

1. zero noise
2. uncorrelated signal plus noise
3. correlated signal and noise

78

Zero Noise. For this case, Glover obtains

$$R_{Wff}(\tau_1, \tau_2) = k_1 \rho_{ff}(\tau_1)\rho_{ff}(\tau_2) + k_2\rho_{ff}(\tau_1 - \tau_2) \tag{120}$$

where

$$k_1 = 1 - \sqrt{\frac{\pi}{2}}\left(\frac{\delta}{\sigma_f} + \frac{\sigma_f}{\delta}\right) e^{\delta^2/2\sigma_f^2} \text{ erfc } \frac{\delta}{\sigma_f\sqrt{2}}$$

$$k_2 = \sqrt{\frac{\pi}{2}}\left(\frac{\sigma_f}{\delta}\right) e^{\delta^2/2\sigma_f^2} \text{ erfc } \frac{\delta}{\sigma_f\sqrt{2}}$$

$\rho_{ff}(\tau)$ = normalized autocorrelation function of the signal (as in Eq 125)

Glover does not present the rather lengthy manipulations required to obtain Eq 120; however, it has been verified independently by the present authors.

The optimum system found from Eq 118 and 120 is the ordinary Wiener system multiplied by a constant, i.e.,

$$k_{opt}(t) = \frac{1}{k} h_W(t) \tag{121}$$

where

$k_{opt}(t)$ = optimum weighting function

$k = B + (1 - B)\int_0^\infty h_W(y)\rho_{ff}(y + \eta)dy$

$B = \dfrac{A}{\dfrac{\delta}{\sigma_f}\left(1 - \dfrac{\delta}{\sigma_f} A\right)}$

$A = \sqrt{\dfrac{\pi}{2}} e^{(1/2)(\delta/\sigma_f)^2} \text{ erfc } \dfrac{\delta}{\sqrt{2}\,\sigma_f}$

$\sigma_f^2$ = variance of signal

$h_W(t)$ = Wiener optimum predictor weighting function

$\rho_{ff}(y + \eta)$ = normalized autocorrelation function

erfc x = the complementary error function

79

Signal Plus Uncorrelated Noise. The general form of the solution for $n(t) \equiv 0$ also applies to the case where $n(t) \neq 0$, providing that $\overline{n(t)} = 0$, and that there is no crosscorrelation between signal and noise. The optimum weighting function is again of the form of Eq 121 where now $h_W(t)$ is the Wiener weighting function for the given signal and noise.

Correlated Signal and Noise. Glover states that attempts to obtain similar solutions both when the noise and true signal are correlated, and when using other types of statistics, lead to the necessity of solving higher-order transcendental equations, which can probably only be solved by recourse to machine computations.

It is a striking fact that for Gaussian inputs and noise, the optimum system is simply the Wiener system multiplied by a constant (Eq 121). Thus, if the optimum Wiener system for a given Gaussian environment is unsatisfactory due to poor transient characteristics, replacing the $\overline{\epsilon^2}$ criterion by Glover's criterion will effect no significant improvement. The corresponding relationship, if any, for non-Gaussian environments has not yet been determined due to the great difficulties associated with general analytic evaluations of Eq 117 and 118.

### Glover's Criterion Applied to Fixed-Order Systems (Phillips-Type Optimization)

General expressions for the optimum P.E. system of fixed-order (Phillips system) will now be derived. (This problem was not considered by Glover.) A typical example is evaluated, and the P.E. performance measure is shown to be even less selective than the minimum mean square error criterion.

Zero Noise. Using the expression for $R_{Wff}(\tau_1, \tau_2)$ given by Eq 120, and inserting it into Eq 115, gives the P.E. as

$$\left[k_1 \rho_{ff}(0)\rho_{ff}(0) + k_2\rho_{ff}(0)\right] - 2\int_0^\infty h(x)\left[\rho_{ff}(-x)\rho_{ff}(0)k_1 + \rho_{ff}(-x)k_2\right]dx$$

$$+ \int_0^\infty \int_0^\infty h(x)h(y)\left[\rho_{ff}(-x)\rho_{ff}(-y)k_1 + \rho_{ff}(y-x)k_2\right]dydx \quad (122)$$

The integrals are merely convolutions evaluated at zero, e.g., the second term is

$$\lim_{t \to 0} 2(k_1 + k_2) \int_0^\infty h(x) \rho_{ff}(t - x) dx \qquad (123)$$

Using the following properties of normalized autocorrelation functions,

$$\rho_{ff}(0) = 1 \qquad (124)$$

$$\rho_{ff}(\tau) = \frac{R_{ff}(\tau)}{\sigma_f^2} \qquad (125)$$

The P.E. becomes

$$\text{P.E.} = \frac{(k_1 + k_2)R_{ff}(0)}{\sigma_f^2} - \frac{2(k_1 + k_2)R_{fc}(0)}{\sigma_f^2} + \frac{k_1 R_{fc}^2(0)}{\sigma_f^4} + \frac{k_2 R_{cc}(0)}{\sigma_f^2} \qquad (126)$$

This can be rewritten by adding and subtracting $\dfrac{k_1 R_{cc}(0)}{\sigma_f^2}$ :

$$\text{P.E.} = \frac{(k_1 + k_2)\left[R_{ff}(0) - 2R_{fc}(0) + R_{cc}(0)\right] + k_1\left\{\dfrac{\left[R_{fc}(0)\right]^2}{\sigma_f^2} - R_{cc}(0)\right\}}{\sigma_f^2} \qquad (127)$$

$$\therefore \quad \sigma_f^2 \text{ P.E.} = (k_1 + k_2)\overline{\epsilon^2} + k_1\left[\frac{R_{fc}^2(0)}{R_{ff}(0)} - R_{cc}(0)\right] \qquad (128)$$

where $\overline{\epsilon^2}$ is the mean square error for this (zero noise) case

It can be seen from Eq 128 that in general the optimum system will not be a constant times the minimum $\overline{\epsilon^2}$ system (unlike the case when the system order and form were left free during the minimization). This can be seen from setting $\partial \text{P.E.}/\partial \alpha_i = 0$ where $\alpha_i$ is a variable parameter of the system.

Equation 128 for P.E. becomes, in this case,

$$\sigma_f^2 \text{ P.E.} = (k_1 + k_2)\overline{\epsilon_x^2} + k_1\left[\frac{R_{fc}^2(0)}{\sigma_f^2} - R_{cc}(0) - R_{c_n c_n}(0)\right] \qquad (129)$$

where $\overline{\epsilon_x^2}$ is the mean square error for the
particular uncorrelated signal and
noise present

$$R_{c_n c_n}(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty}|H(j\omega)|^2\Phi_{nn}(\omega)d\omega$$

Again it can be seen that in general, minimizing P.E. will not lead to the same parameters of $H(j\omega)$ as minimizing $\overline{\epsilon^2}$.

Signal Plus Correlated Noise. The expression for P.E. for this case will be similar to those of the previous cases. The result will not be detailed here because the expressions are very lengthy. However, the _form_ of the P.E. expressions has been studied, and, as in the case of signal plus uncorrelated noise, it is such that the optimum P.E. system is in general not simply related to the optimum $\overline{\epsilon^2}$ system.

## Example of Phillips-Type Optimization Using Glover's Criterion

As an example of Phillips-type optimization, consider a second-order system,
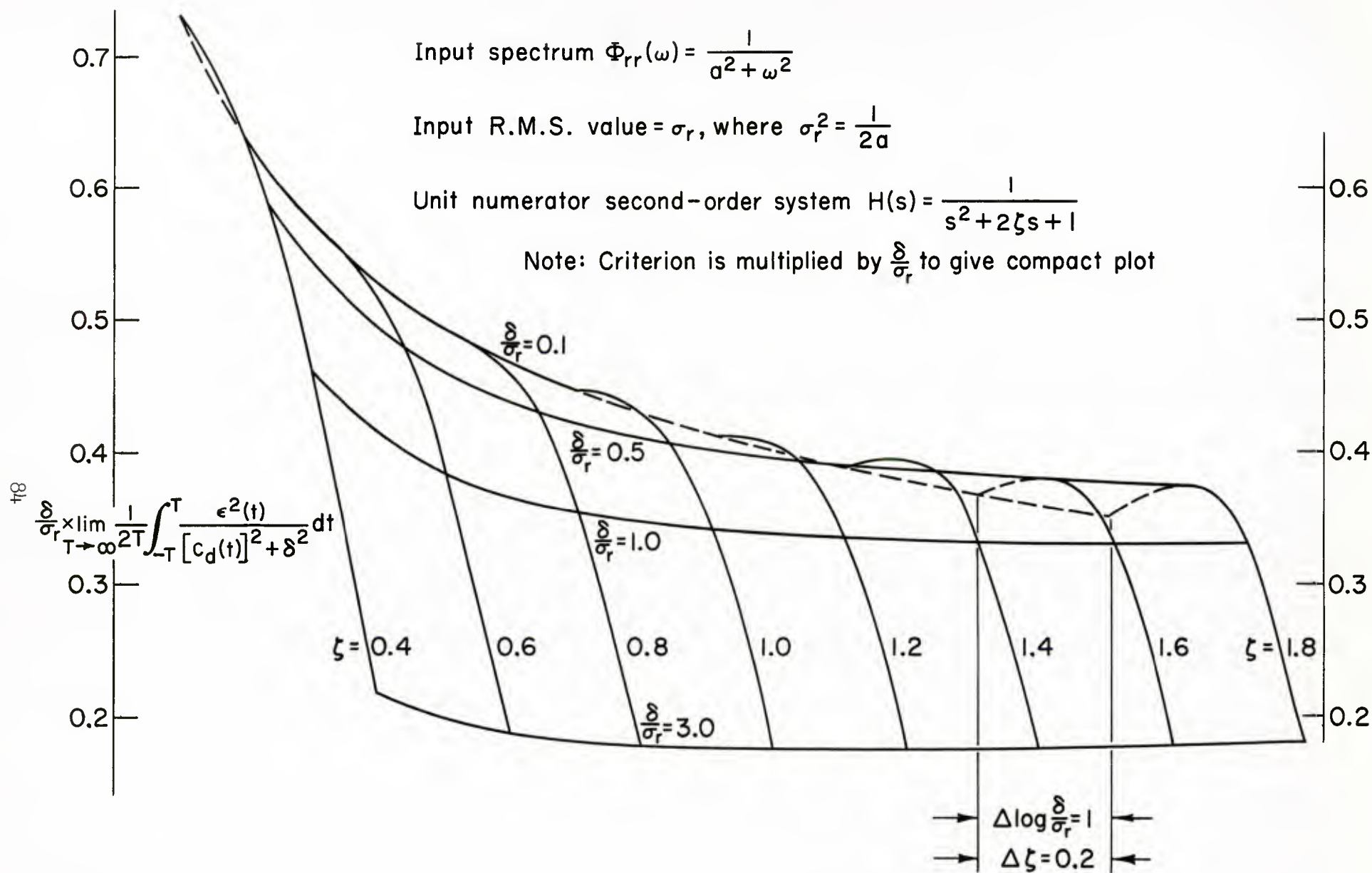
$$H(s) = \frac{1}{s^2 + 2\zeta s + 1} \qquad (130)$$

with a signal input spectrum

$$\Phi_{ff}(\omega) = \frac{1}{a^2 + \omega^2} \qquad (131)$$

$$\Phi_{nn}(\omega) = 0$$

The variation of Glover's performance measure for this system is illustrated in Fig. 22. In Fig. 23, Glover's criterion is directly compared with the criterion

of minimum $\overline{\epsilon^2}$. This comparison has been effected by combining the representative $\delta/\sigma_r = 1.0$ graph of Fig. 22 (which corresponds to $\delta$ = rms input amplitude) with the a = 0.25 graph of Fig. 5, which corresponds to an input cutoff frequency of 0.25 of the system undamped natural frequency. Figure 23 shows that Glover's criterion is less selective than the minimum $\overline{\epsilon^2}$ criterion. But the selectivity of the $\overline{\epsilon^2}$ criterion is only barely acceptable. Therefore, Glover's criterion—due to its poor selectivity—must be regarded as unsatisfactory.

Figure 22. $\frac{\delta}{\sigma_r}$ x Glover's Criterion: Carpet of $\frac{\delta}{\sigma_r} \lim\limits_{T\to\infty} \frac{1}{2T} \int_{-T}^{+T} \frac{\epsilon^2(t)}{[c_d(t)]^2 + \delta^2}\, dt$ vs $\zeta$ and $\frac{\delta}{\sigma_r}$
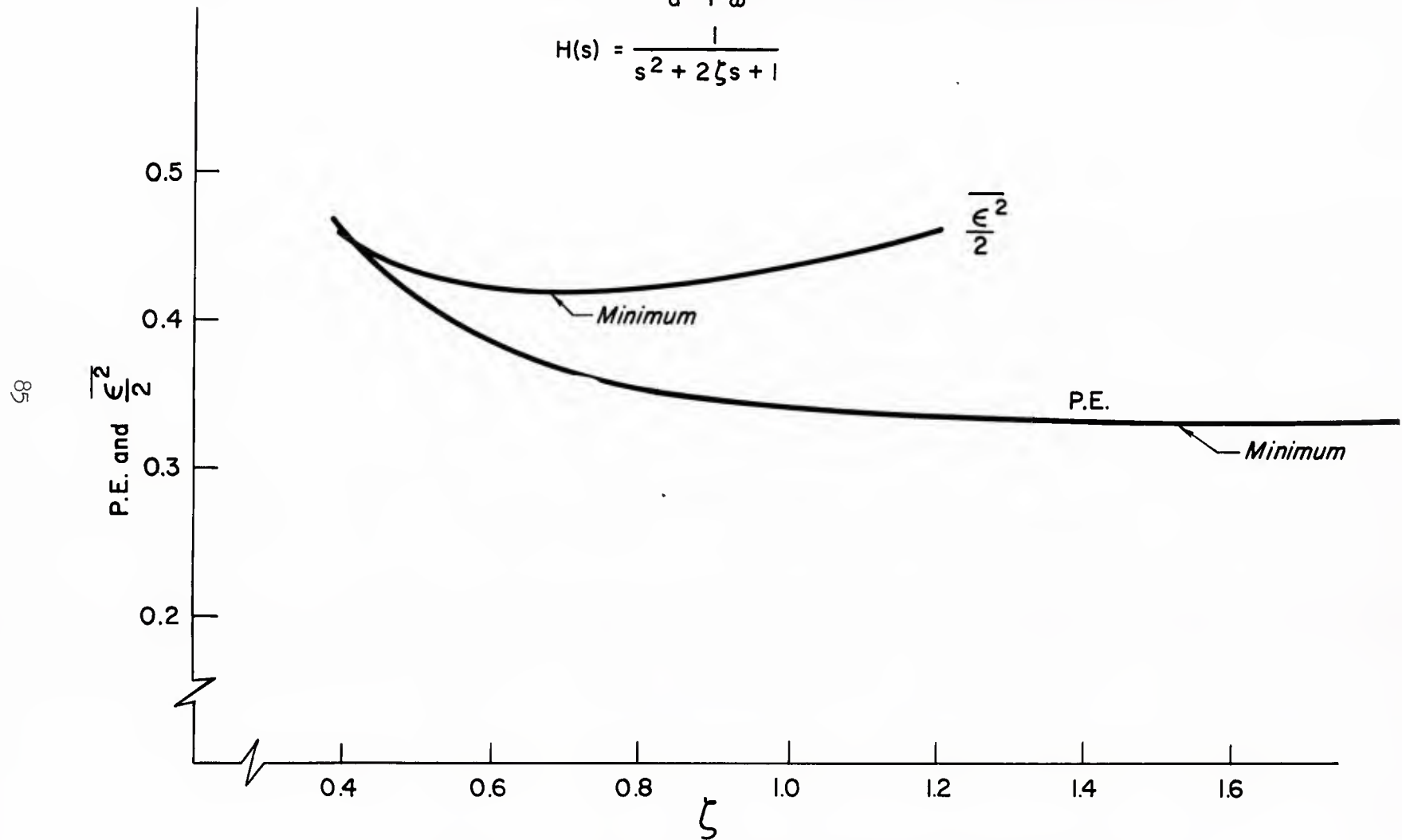
Figure 23. Comparison of Glover's (P.E.) Criterion and Mean Square Error of Second-Order Unit Numerator System

CHAPTER IV

THE GENERALIZED ERROR FUNCTION

This chapter is primarily concerned with the generalized error function (G.E.F.) which, for stationary random inputs (zero noise), is defined as

$$\text{G.E.F.} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} \left[ \epsilon(t, \tau) \right]^2 dt \equiv \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} \left[ r(t - \tau) - c(t) \right]^2 dt$$

(132)

The G.E.F. has a compatible form for deterministic inputs, which Ref. 31 denotes by $E_t$:

$$E_t = \int_0^\infty \left[ \epsilon(t, \tau) \right]^2 dt \equiv \int_0^\infty \left[ r(t - \tau) - c(t) \right]^2 dt$$

(133)

As defined in the Introduction, compatibility implies that $E_t$ and G.E.F. will be numerically equal for any given linear constant-coefficient system, provided that the appropriate deterministic input used in computing $E_t$ is the transient analog of the random input. It has already been shown (p. 14) that $\left[ \text{G.E.F.} \right]_{\tau=0}$ and $\left[ E_t \right]_{\tau=0}$ are compatible for the case of a unit numerator second-order system when the random input has the power spectrum $\Phi_{rr}(\omega) = 1/(\omega^2 + a^2)$ and the deterministic input is $e^{-at}$. In this chapter it will be shown that the G.E.F. and $E_t$ are compatible for all linear constant-coefficient systems, i.e., the transient analog of a random input does not depend upon the system to which it is applied. This is demonstrated in Section A below by expansion of the analysis of Benedict and Rideout (Ref. 29), who considered only the case of zero noise. Section B derives general transient analogs for both uncorrelated and correlated signal and noise inputs. These results are used extensively in the third part of the chapter to calculate and form an assessment of the G.E.F. Since, for $\tau = 0$, G.E.F. = $\overline{\epsilon^2}$, the transient analog concepts apply equally well to the optimization of systems using the mean square error criterion.

A. TRANSIENT ANALOGS (ZERO NOISE)

For the G.E.F. and $E_t$ performance measures to be compatible, the transient analog input must have an energy spectral density that is the same function of

frequency (except for the units) as the power spectral density of the random input signal. This relationship was first noted by Schultz and Rideout (Ref. 28). A proof is given below.

The Generalized Error Function for either random or deterministic inputs can be thought of as the mean square error where the error is defined in Fig. 24.
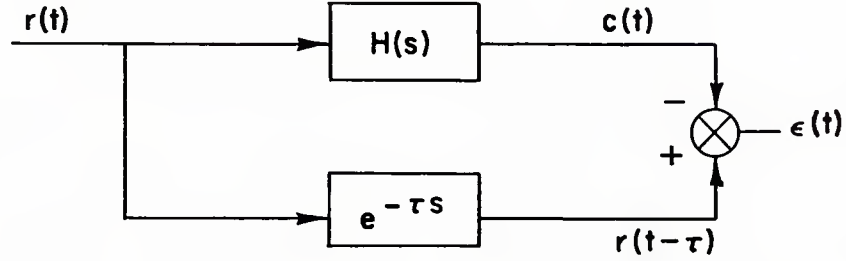


Figure 24. Block Diagram for Computation of the G.E.F.

For a transient deterministic input, $r_t$, commencing at $t = 0$, $E_t$ (Eq 133) can be expressed as

$$E_t(\tau) = \int_0^\infty \left[r_t(t - \tau)\right]^2 dt - 2\int_0^\infty r_t(t - \tau)c_t(t)dt + \int_0^\infty \left[c_t(t)\right]^2 dt \tag{134}$$

The corresponding expression for the G.E.F. (Eq 132) for a stationary random input, r, with output c, can be written in terms of auto- and crosscorrelation functions as

$$\text{G.E.F.} = R_{rr}(0) - 2R_{rc}(\tau) + R_{cc}(0) \tag{135}$$

It will now be shown that Eq 134 and 135 are identical term by term, if the deterministic input, $r_t(t)$, is the transient analog of the random signal, $r(t)$ (the energy spectral density of $r_t(t)$ being equal to the power spectral density of $r(t)$). Taking the last term first, by Parseval's theorem,

$$\int_0^\infty \left[c_t(t)\right]^2 dt = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} C_t(s)C_t(-s)ds = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} H(s)R_t(s)H(-s)R_t(-s)ds \tag{136}$$

where $H(s)$ = the system transfer function in Fig. 24

$R_t(s) = \mathcal{L}\left[r_t(t)\right]$

88

Substituting $s = j\omega$, Eq 136 can be written in terms of the input energy spectral density as

$$\int_0^\infty [c_t(t)]^2 dt = \frac{1}{2\pi} \int_{-\infty}^\infty H(j\omega)H(-j\omega) S_{r_t r_t}(\omega) d\omega \qquad (137)$$

where $S_{r_t r_t}(\omega) \equiv R_t(j\omega)R_t(-j\omega)$, the transient input energy spectral density (p. 102 of Ref. 50)

The last term of the G.E.F. (Eq 135) can be written as

$$R_{cc}(0) = \frac{1}{2\pi} \int_{-\infty}^\infty \Phi_{cc}(\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^\infty H(j\omega)H(-j\omega)\Phi_{rr}(\omega) d\omega \qquad (138)$$

and if

$$\Phi_{rr}(\omega) \equiv S_{r_t r_t}(\omega) \qquad (139)$$

then from Eq 137 and 138 it can be seen that

$$R_{cc}(0) = \int_0^\infty [c_t(t)]^2 dt \qquad (140)$$

Similarly, the first terms of Eq 134 and 135 are equal.

$$\int_0^\infty [r_t(t - \tau)]^2 dt = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} [e^{-\tau s}R_t(s)][e^{\tau s}R_t(-s)] ds \qquad (141)$$

$$= \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} R_t(s)R_t(-s) ds \qquad (142)$$

$$= \frac{1}{2\pi} \int_{-\infty}^\infty S_{r_t r_t}(\omega) d\omega \qquad (143)$$

and, if the energy spectrum of $r_t(t)$ is the same function of $\omega$ as the power spectrum of $r(t)$,

89

$$\int_0^\infty [r_t(t - \tau)]^2 dt = \frac{1}{2\pi} \int_{-\infty}^\infty \Phi_{rr}(\omega) d\omega \qquad (144)$$

The middle term of Eq 134 can be shown to be equal to the middle term of Eq 135 by using the following version of Parseval's theorem (Ref. 6, p. 43):

$$\int_0^\infty x_1(t)x_2(t) dt = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} X_1(-s)X_2(s) ds \qquad (145)$$

where $X_1(s) = \mathcal{F}[x_1(t)]$
$X_2(s) = \mathcal{F}[x_2(t)]$

Substituting $r_t(t - \tau)$ for $x_1(t)$ and $c_t(t)$ for $x_2(t)$ yields the middle term of Eq 134 as

$$\int_0^\infty r_t(t - \tau)c_t(t) dt = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} e^{\tau s} R_t(-s)C_t(s) ds \qquad (146)$$

where $X_1(s) = R_t(s)e^{-\tau s}$
$X_2(s) = C_t(s)$

Equation 146 can be written as

$$\int_0^\infty r_t(t - \tau)c_t(t) dt = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} H(s)R_t(s)R_t(-s)e^{\tau s} ds \qquad (147)$$

Substituting $s = j\omega$, $\qquad = \frac{1}{2\pi} \int_{-\infty}^\infty H(j\omega) S_{r_t r_t}(\omega) e^{j\omega\tau} d\omega \qquad (148)$

Now the middle term of Eq 135 is (by definition)

$$R_{rc}(\tau) = \frac{1}{2\pi} \int_{-\infty}^\infty \Phi_{rc}(j\omega) e^{j\omega\tau} d\omega \qquad (149)$$

and since $\Phi_{rc}(j\omega) = H(j\omega)\Phi_{rr}(\omega)$, Eq 149 becomes

90

$$R_{rc}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(j\omega)\Phi_{rr}(\omega)e^{j\omega\tau}d\omega \tag{150}$$

If the energy spectrum of $r_t(t)$ is the same function of frequency as the power spectrum of $r(t)$, then Eq 150 and 148 are equal, i.e.,

$$\int_0^{\infty} r_t(t - \tau)c_t(t)dt = R_{rc}(\tau) \tag{151}$$

Equations 134 and 135 are then identical term by term for all H(s) if $r_t(t)$ is the transient analog of $r(t)$, i.e., if $\Phi_{rr}(\omega) = \mathcal{S}_{r_t r_t}(\omega)$.

A transient signal whose energy spectral density is the same function of frequency as the power spectral density of any stationary random signal can be obtained as indicated by Fig. 25 where the white noise generator has unit power per cps.



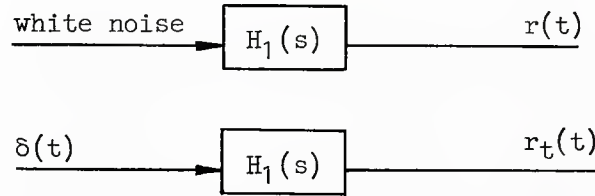Figure 25. Generation of Transient Analog, $r_t(t)$

The random signal spectrum is

$$\Phi_{rr}(\omega) = H_1(j\omega)H_1(-j\omega) \tag{152}$$

while the transient signal is

$$\begin{aligned} r_t(t) &= h_1(t) \\ R_t(s) &= H_1(s) \end{aligned} \tag{153}$$

and therefore its energy spectral density is

$$\mathcal{S}_{r_t r_t}(\omega) = R_t(j\omega)R_t(-j\omega) = H_1(j\omega)H_1(-j\omega) \tag{154}$$

91

Before proceeding to describe transient analogs for signal plus noise inputs (correlated or uncorrelated), two minor points must be noted. Throughout the above analysis it has been tacitly assumed that the value of $\tau$ is fixed. In fact, the selection of $\tau$ is not a simple matter. The reader may also have observed that for an integrated white noise input (which has a step as its transient analog), Eq 134 cannot be evaluated directly because the integrals on the right fail to converge. This obstacle can be easily circumvented. Discussion of both these details is delayed until later in this chapter to avoid too great a digression from the present topic of transient analogs for signal and noise.

B.  TRANSIENT ANALOGS FOR UNCORRELATED SIGNAL AND NOISE

The G.E.F. for a system subjected to uncorrelated signal and noise can be obtained very simply by applying transient analogs of the signal and of the noise at widely separated times, as will now be shown.
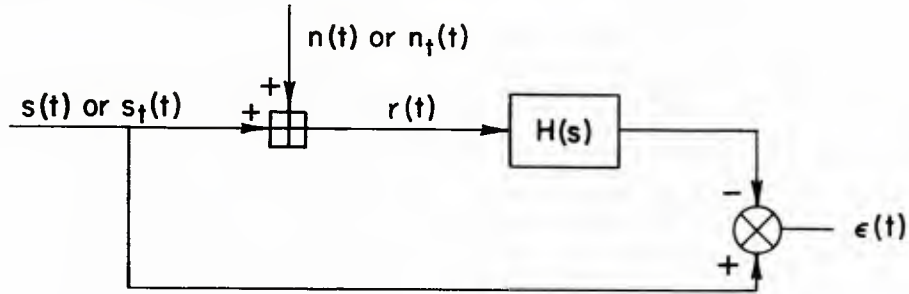


Figure 26.  Block Diagram for Uncorrelated Signal and Noise Inputs

For the system illustrated in Fig. 26, the error power spectrum is (from p. 50)

$$\Phi_{\epsilon\epsilon}(\omega) \quad = \quad |1 - H(j\omega)|^2 \Phi_{ss}(\omega) \; + \; |H(j\omega)|^2 \Phi_{nn}(\omega) \tag{155}$$

Zero prediction or lag ($\tau = 0$) will be assumed to keep the analysis brief, although there is no essential difficulty in extending it to cover the case of $\tau \neq 0$. With $\tau = 0$,

$$\text{G.E.F.} \quad = \quad \overline{\epsilon^2} \quad = \quad \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_{\epsilon\epsilon}(\omega)\,d\omega \tag{156}$$

The transient analog of the input will now be shown to consist of the transient analog of the signal, $s_t(t)$, plus that of the noise, $n_t(t)$, where $n_t(t)$ is applied

92

to the system only after the error transients produced by $s_t(t)$ have become effectively zero. Therefore, the $IE^2$ for these transient signals is

$$IE^2 = \int_0^\infty \left[\epsilon_1(t) + \epsilon_2(t - T)\right]^2 dt \tag{157}$$

where    $\epsilon_1(t)$    is the error due to the signal transient analog input, $s_t(t)$ applied at $t = 0$

$\epsilon_2(t - T)$    is the error due to the noise transient analog input, $n_t(t)$ applied at $t = T$

$T$    is much longer than the system settling time

Since $\epsilon_2(t - T)$ only occurs long after the initiation of $\epsilon_1(t)$, there will be no crossproduct; therefore,

$$IE^2 = \int_0^T \left[\epsilon_1(t)\right]^2 dt + \int_T^\infty \left[\epsilon_2(t - T)\right]^2 dt \tag{158}$$

The first and second integrals are approximately $IE_1^2$ and $IE_2^2$; therefore,

$$IE^2 \doteq \int_0^\infty \epsilon_1^2 dt + \int_0^\infty \epsilon_2^2 dt \tag{159}$$

and using Parseval's theorem, this can be written as (defining $E_i(s) = \mathcal{L}|\epsilon_i(t)|$)

$$IE^2 = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} E_1(s)E_1(-s)ds + \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} E_2(s)E_2(-s)ds \tag{160}$$

$E_1(s)$ and $E_2(s)$ can be found from the definitions below Eq 157, and Fig. 26. $|E_1(s)|^2$ and $|E_2(s)|^2$ can then be expressed in terms of the signal or noise energy spectral densities.

$$\mathcal{S}_{s_t s_t}(s) = S(s)S(-s) \quad \text{where} \quad S(s) = \mathcal{L}\left[s_t(t)\right]$$
$$\mathcal{S}_{n_t n_t}(\omega) = N(s)N(-s) \quad \text{where} \quad N(s) = \mathcal{L}\left[n_t(t)\right] \tag{161}$$

Equation 160 becomes

$$IE^2 = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \left[|1 - H|^2 \mathcal{S}_{s_t s_t} + |H|^2 \mathcal{S}_{n_t n_t}\right] ds \tag{162}$$

93

Since $\Phi_{ss} = S_{s_t s_t}$ and $\Phi_{nn} = S_{n_t n_t}$, the integrand in Eq 162 equals the right side of Eq 155 term by term, and therefore $IE^2 \equiv \overline{\epsilon^2}$.

Example of Transient Analogs for Uncorrelated Signal and Noise. A random signal having the power spectrum $\Phi_{ss}(\omega) = 1/(a^2 + \omega^2)$ has a transient analog $s_t(t) = e^{-at}$ [because the energy spectral density of this transient is $|S_t(s)S_t(-s)|_{s=j\omega} = 1/(a^2 + \omega^2)$]. Thus, the $\overline{\epsilon^2}$ produced by a stationary random signal, $\Phi_{ss}(\omega) = 1/(a^2 + \omega^2)$, in the presence of white noise, $\Phi_{nn}(\omega) = 1$, can be obtained by applying the inputs, $s_t(t) = e^{-at}$, and a delayed unit impulse, $n_t(t) = \delta(t - T)$, and integrating the error squared of the transient response. The shaded area of Fig. 27 indicates the parts of the error response that contribute to the total $IE^2$.
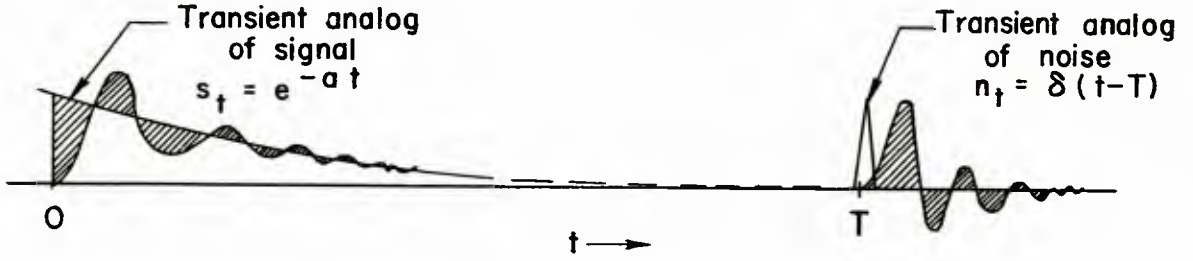


Figure 27. Transient Analogs for Signal and Noise

## C. TRANSIENT ANALOGS FOR CORRELATED SIGNAL AND NOISE

The transient analog for correlated signal and noise is the transient analog for uncorrelated signal and noise plus an additional transient input to a modified system that simulates the effect of the correlation. This modified system can be varied independently of the signal and noise spectra to determine the effect of the crosscorrelation on the G.E.F.

For the system of Fig. 26 with correlated signal and noise (Ref. 50, p. 239), the error power spectral density is

$$\Phi_{\epsilon\epsilon} = [1 - H]^2 \Phi_{ss} + [H]^2 \Phi_{nn} - (1 - H^*)H\Phi_{sn} - (1 - H)H^*\Phi_{ns} \qquad (163)$$

94

where the following abbreviated notation is used to keep the expressions reasonably concise:

$$|H|^2 \equiv HH^*$$

$$H \equiv H(j\omega)$$

$$H^*(j\omega) \equiv H(-j\omega)$$

$$\Phi_{sn} \equiv \Phi_{sn}(j\omega)$$

$$\Phi_{ns} \equiv \Phi_{ns}(j\omega)$$

$$\Phi_{ss} \equiv \Phi_{ss}(\omega)$$

$$\Phi_{nn} \equiv \Phi_{nn}(\omega)$$

(164)

and the mean square error is

$$\overline{\epsilon^2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_{\epsilon\epsilon}(\omega) d\omega \tag{165}$$

The first two terms of Eq 163 do not involve the crosscorrelation, and their transient analogs can be obtained by the procedure that is described in Section B. The contribution of the last two terms of Eq 163 to $\overline{\epsilon^2}$ can be expressed as follows:

$$\overline{\Delta\epsilon^2} = -\frac{1}{2\pi} \int_{-\infty}^{\infty} (1 - H^*)H\Phi_{sn} d\omega - \frac{1}{2\pi} \int_{-\infty}^{\infty} (1 - H)H^*\Phi_{ns} d\omega \tag{166}$$

It is assumed that $\Phi_{sn}(j\omega) = \Phi_{ns}(-j\omega)$ is a rational function of $j\omega$, i.e.,

$$\Phi_{sn}(j\omega) = H_a(j\omega)H_b(-j\omega) = H_a H_b^* \tag{167}$$

where    $H_a$   is a system transfer function composed of the left-half-s-plane poles and zeros of   $\Phi_{sn}(j\omega)$

$H_b^*$   is similarly composed of the right-half-s-plane poles and zeros of   $\Phi_{sn}(j\omega)$

Inserting Eq 167 into Eq 166 and substituting $s = j\omega$ yields (after rearranging terms)

$$\overline{\Delta\epsilon^2} = -\frac{1}{2\pi j} \int_{-j\infty}^{j\infty} (1 - H^*)H_b^* H_a H ds - \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} H^* H_a^* H_b (1 - H) ds \tag{168}$$

Using a version of Parseval's theorem (Ref. 6, p. 43),

95

$$\int_0^\infty x_1(t)x_2(t)dt = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} X_1(-s)X_2(s)ds \qquad (169)$$

$$\text{where} \quad X_1(s) = \mathcal{L}\left[x_1(t)\right]$$
$$X_2(s) = \mathcal{L}\left[x_2(t)\right]$$

it can be shown that $\overline{\Delta\epsilon^2}$ is twice either term in Eq 168:

$$\overline{\Delta\epsilon^2} = -2\left[\frac{1}{2\pi j} \int_{-j\infty}^{j\infty} (1 - H^*)H_b^* H_a H ds\right] \qquad (170)$$

This can be seen from Eq 168 and 169. In Eq 168 the integrands are complex conjugates. Identifying the asterisked quantities $(1 - H^*)$ and $H_b^*$ in the first integrand of Eq 168 with $X_1(-s)$ and the remaining quantities with $X_2(s)$ reveals that the first integral of Eq 168 is equal to the left side of Eq 169. A similar identification can be made for the second integral of Eq 168. Therefore, these integrals are equal. Equation 170 for $\overline{\Delta\epsilon^2}$ is now in a form suitable for computation by means of operations upon transient quantities. A circuit suitable for the synthesis of $\overline{\Delta\epsilon^2}$ is illustrated in Fig. 28, where transient inputs $\mathcal{L}^{-1}H_a(s)$ and $\mathcal{L}^{-1}H_b(s)$ are generated simultaneously and fed through systems having transfer functions $H(s)$ and $1 - H(s)$. The resulting signals, $\epsilon_a(t)$ and $\epsilon_b(t)$, are multiplied and integrated to give $\overline{\Delta\epsilon^2}$. This can be seen by inspection of Fig. 28 and use of Eq 169 to yield Eq 170. Note that since $H_a(s)$ and $H_b(s)$ both have left-half-plane poles only, there is no problem with regard to stability or physical realizability of this circuit.
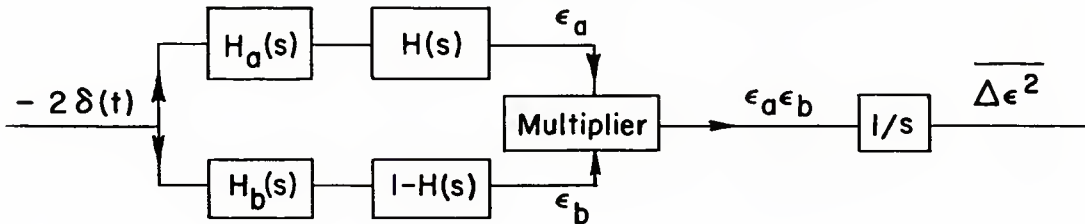


Figure 28. Circuit for Generating Contribution to $\overline{\epsilon^2}$ Due to Crosscorrelation of Signal and Noise

96

Etkin (Eq 2.17 through 2.19 of Ref. 44) derives an incorrect general form for the transient analog of a system subjected to two concurrent stationary random inputs having power spectra $\Phi_1(\omega)$ and $\Phi_2(\omega)$, and a cross spectrum $\Phi_{12}(j\omega)$. (These could, of course, represent correlated signal and noise.) His analysis is valid only when

$$R_{t_s}(j\omega) R^*_{t_n}(j\omega) \quad = \quad 2\pi \Phi_{ns}(j\omega) \tag{171}$$

There is no reason why Eq 171 (which corresponds to Eq 2.19 of Ref. 44) should be satisfied in general, since it says that the cross-spectrum between signal and noise <u>always</u> has the same left-half-plane poles and zeros as the noise spectrum and the same right-half-plane poles and zeros as the signal spectrum. It is, however, physically possible, and indeed probable, that the cross-spectrum has poles and zeros not contained in either the signal or the noise spectra.

D.   EVALUATION OF THE G.E.F.

Having discussed the transient analogs associated with the G.E.F. (Eq 132), the measure itself will now be examined.  Its compatible criterion (Eq 133)

$$\text{minimum } E_t(\tau) \quad = \quad \text{minimum} \int_0^\infty [r(t - \tau) - c(t)]^2 dt \tag{172}$$

was originally considered by Aigrain and Williams (Ref. 52) for the optimization of amplifiers subject to step inputs.  Schultz (Ref. 48) and Schultz and Rideout (Ref. 28) continued the investigation of $E_t(t)$ and linked it directly to the G.E.F.  The first use of the G.E.F. for random inputs was by Lee and Wiesner (Ref. 54), further discussion being given by Spooner and Rideout (Ref. 30). Minimum G.E.F. can also be regarded as an alternate form of the Wiener optimum predicting ($\eta > 0$) or lagging ($\eta < 0$) criterion (see p. 20 ):

$$\text{minimum}_{\substack{\text{predicting} \\ \text{or lagging}}} \overline{\epsilon^2} \quad = \quad \text{minimum} \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} [r(t - \eta) - c(t)]^2 dt \tag{173}$$

The G.E.F. is thus very similar to the $\overline{\epsilon^2}$ for random inputs, and $E_t$ is closely related to $IE^2$ for deterministic inputs.  Much of the following section can be

97

interpreted as a discussion of $\overline{\epsilon^2}$ for cases where the predicting or lagging time, $\eta$, equals $\tau$, a time constant which can be chosen freely.

As has been shown in Ref. 1 and 34, the $IE^2$ performance measure exhibits poor selectivity for step inputs, i.e., the $IE^2$ of the optimum system is only a little less than that of a wide range of off-optimum systems. Fig. 5 demonstrates that the minimum $\overline{\epsilon^2}$ criterion is equally unselective for a typical random input. It has been suggested that the G.E.F. might achieve greater selectivity because it compares the output with the delayed input, and the system is therefore not penalized for unavoidable initial errors. However, the principal conclusions of this section are

1.  the G.E.F. is actually little more selective than the mean square error criterion

2.  difficulties frequently occur in selecting the value of $\tau$ to be used

3.  the G.E.F. does have some merit in that it possesses compatible forms for both stationary random and deterministic inputs, and largely because of this advantage, studies of modified forms of G.E.F. may be worthwhile

To use the generalized error function, it must be put into concrete form by suitable choice of $\tau$. This problem will now be discussed.

Selection of $\tau$

This choice could be related to some of the parameters of the system, but perhaps the simplest alternative is to make the time delay, $\tau$, equal to the delay time, $t_d$.

$$c(t_d) = \frac{1}{2} c(\infty) \tag{174}$$

> where $c(\infty)$ is the steady state value of the system step response, and $t_d$ is defined as the time to achieve 50 percent of the final value of the step (indicial) response

This choice also minimizes the $E_t(\tau)$ for a step input, as will now be shown.

$$E_t(\tau) = \int_0^\infty \left[ r(t - \tau) - c(t) \right]^2 dt, \text{ for } r(t) = u(t) \tag{175}$$

$$= \int_0^\infty \left( \left[ r(t - \tau) - r(t) \right] - \left[ c(t) - r(t) \right] \right)^2 dt \tag{176}$$

$$= \int_0^\infty \left[ \epsilon(t) \right]^2 dt + \int_0^\tau \left[ r(\infty) \right]^2 dt - 2 \left[ r(\infty) \right] \int_0^\tau \epsilon(t) dt \tag{177}$$

Differentiating with respect to $\tau$,

$$\frac{dE_t(\tau)}{d\tau} = \left[ r(\infty) \right]^2 - 2 \left[ r(\infty) \right] \epsilon(\tau) \tag{178}$$

$\therefore$ For $E_t(\tau)$ to be a minimum,

$$\epsilon(\tau) = \frac{1}{2} r(\infty) \tag{179}$$

$E_t(\tau)$ therefore has a minimum when $\tau$ is equal to $t_d$, the delay time. (In Ref. 28 this fact is demonstrated by use of a rather cumbersome geometric argument.) Note that the determination of $\tau = t_d$ is generally impossible by analytic means, since $t_d$ is the solution of a transcendental equation, e.g.; for a third-order response, Eq 179 becomes

$$Ae^{-\gamma_1 t_d} + Be^{-\gamma_2 t_d} \sin (\beta t_d + \psi) = \frac{1}{2} \tag{180}$$

$\tau$ must therefore be chosen by solving Eq 180 approximately, or by using the standard approximations given in Ref. 1.

The above discussion has implications for random as well as step inputs. Using the transient analog concept, the G.E.F. for any random input can be minimized by choosing $\tau$ to minimize $E_t$ for the analogous transient input. No practical random input yields a step for its transient analog, because, as shown in Eq 144, this would imply infinite input power. However, if the input bandwidth significantly exceeds the system bandwidth, the analogous transient

input will have time constants much greater than those of the system, and it will often be possible to approximate the actual transient analog signal by a step, as far as the choice of $\tau$ is concerned, i.e., $\tau \doteq (t_d)_{\text{step input}}$, for minimum G.E.F. For input bandwidths less than the system bandwidth, this artifice will fail and the selection of $\tau$ for minimum G.E.F. must be made on a trial-and-error basis. Fig. 29 illustrates the generalized error function for a second-order system,

$$H(s) \;=\; \frac{1}{s^2 + 2\zeta s + 1} \tag{181}$$

having an input signal power spectrum $\Phi_{rr}(\omega) = 1/(1 + \omega^2 T_1^2)$ with zero noise and $T_1 = 1$. In this case the value of $\tau$ at which the minimum G.E.F. occurs is hardly affected by $\zeta$, but there does not appear to be any reason why this should generally be true. (Note that although Ref. 30 contains several misprints, the data graphed in Fig. 29 have been checked and found to be correct.) Even for this simple case, the evaluation of the G.E.F. is fairly tedious, and in fact a digital computer was employed in Ref. 30. Figure 29 (Fig. 4 of Ref. 30) shows that the selectivity of the G.E.F. with respect to $\zeta$ is poor. For $\tau = 1.15$, changing $\zeta$ from 0.5 to 0.7 raises $E_t$ by only 4 percent, approximately. $E_t(\tau)$ is minimized by $\zeta = 0.5$ at $\tau = 1.15$. G.E.F. thus gives about the same validity and selectivity as the simpler $\overline{\epsilon^2}$ criterion, and appears unpromising as a performance criterion. Some calculations have been performed for other values of $T_1$ which do not change this conclusion. On p. 323 of Ref. 30, alternate forms of G.E.F. are suggested, including one with multiple delays:

$$E(\tau, a, \tau_1, \tau_2) \;=\; r(t) - ac(t + \tau_1) - (1 - a)c(t + \tau_2) \tag{182}$$

where a is a constant, and $\tau_1$ and $\tau_2$ are time constants

It is possible that such a G.E.F. employing multiple delays might achieve adequate selectivity. Some further investigation along these lines might be worthwhile, particularly if digital computing aids are available. The problem of selecting the optimum $\tau$ would then become less acute, because $E(\tau)$ could be swiftly calculated for a wide range of $\tau$.
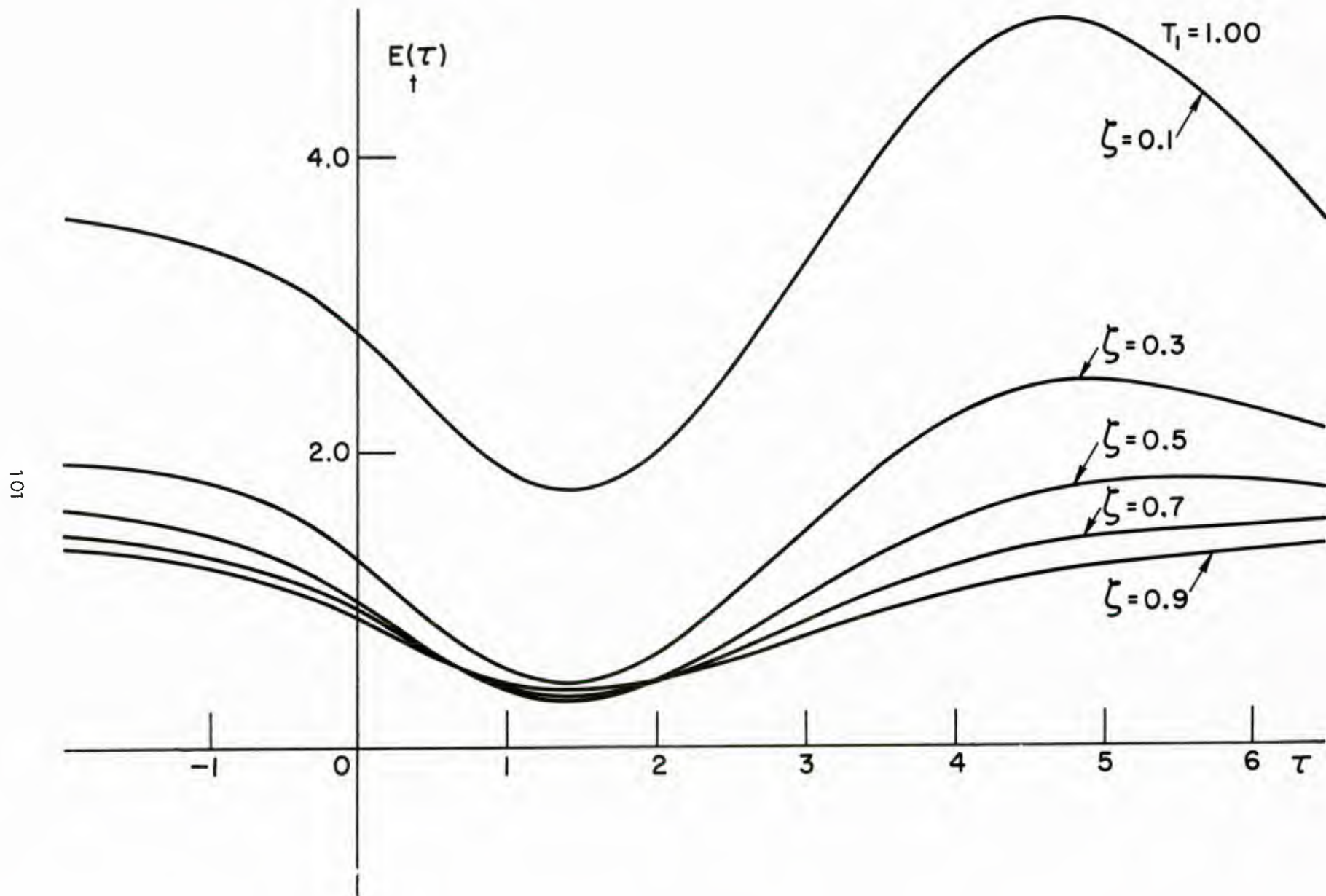
Figure 29. Generalized Error Function for Second-Order Unit Numerator System $H(s) = \dfrac{1}{s^2 + 2\zeta s + 1}$ with Input Power Spectrum, $\Phi_{rr} = \dfrac{1}{1 + T_1 \omega^2}$, with $T_1 = 1$

A Note on the Evaluation of G.E.F. for an Integrated White Noise Input

The transient analog of an integrated white noise input, $\Phi_{rr}(\omega) = 1/\omega^2$, is a step, $R(s) = 1/s$. As noted previously, for such inputs it is difficult to evaluate the G.E.F. using Eq 134, because the integrals fail to converge. In order to demonstrate how this disadvantage can be circumvented, the G.E.F. for a second-order unit numerator system forced by an integrated white noise input will now be evaluated.

The system transfer function is

$$H(s) \quad = \quad \frac{1}{s^2 + 2\zeta s + 1} \tag{183}$$

Evaluating $E_t$, for a unit step input, instead of the G.E.F. for the actual integrated white noise input,

$$E_t(\tau) \quad = \quad \int_0^\infty \left[r(t - \tau) - c(t)\right]^2 dt \tag{184}$$

Adding and subtracting $r(t)$ to the kernel of the integrand overcomes the convergence difficulty.

$$E_t(\tau) \quad = \quad \int_0^\infty \left[r(t) - c(t) + r(t - \tau) - r(t)\right]^2 dt \tag{185}$$

$$= \quad \int_0^\infty \epsilon^2(t) + 2\epsilon(t)\left[r(t - \tau) - r(t)\right] + \left[r(t - \tau) - r(t)\right]^2 dt \tag{186}$$

$$E_t(\tau) \quad = \quad IE^2 + 2\int_0^\infty \epsilon(t)\left[u(t - \tau) - u(t)\right] dt + \int_0^\infty \left[u(t - \tau) - u(t)\right]^2 dt \tag{187}$$

$$= \quad IE^2 - 2\int_0^\tau \epsilon(t) dt + \tau \tag{188}$$

$$\text{where} \quad IE^2 \quad = \quad \int_0^\infty \left[\epsilon(t)\right]^2 dt$$

102

But
$$\epsilon(t) = \mathcal{L}^{-1}\left[\frac{s^2 + 2\zeta s}{s^2 + 2\zeta s + 1}\right] = \frac{e^{-\zeta t}}{\beta} \sin(\beta t + \psi) \tag{189}$$

where $\beta = 1 - \zeta^2$

$\psi = \sin^{-1}\beta$

$$\therefore \quad E_t(\tau) = IE + \tau - 4\zeta + A \tag{190}$$

where $A = \dfrac{2e^{-\tau\zeta}}{\beta}\left[(\zeta^2 - \beta^2)\sin\beta\tau + 2\zeta\beta\cos\beta\tau\right]$

The choice of $\zeta$ for minimum $E_t(\tau)$ requires that the equation $\left[\partial E_t(\tau)/\partial\zeta\right]_{\tau=t_d} = 0$ be solved. This equation is transcendental and no analytic solution exists. The results shown in Fig. 30 (Fig. 4 of Ref. 28) were obtained by digital computation. The selectivity is little better than that of Fig. 29. This result supports the general assertion made above that the G.E.F. (in its present form) has poor selectivity.
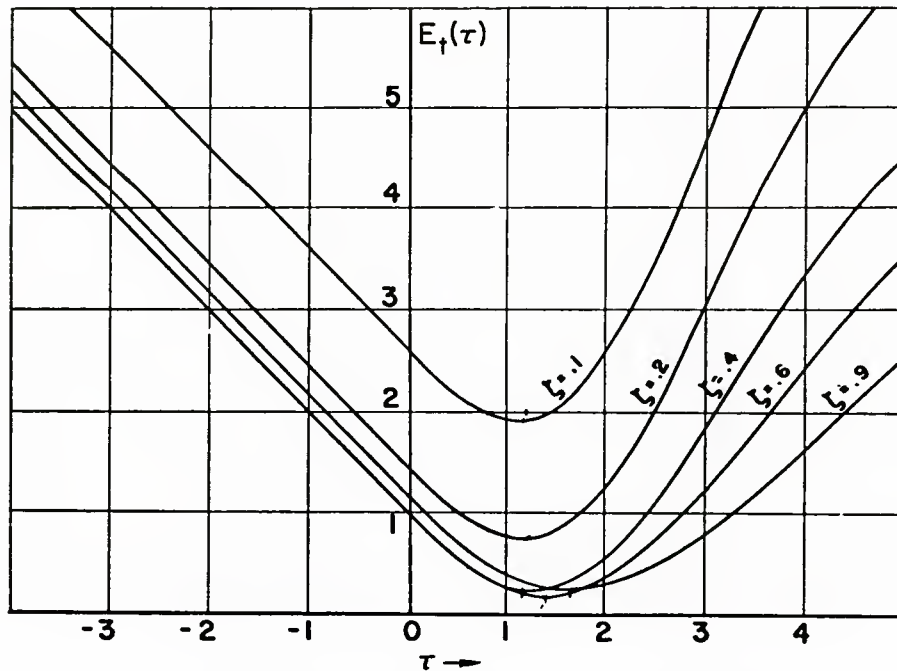


Figure 30. $E_t(\tau)$ vs $\tau$ for a Second-Order Unit Numerator System $H(s) = \dfrac{1}{s^2 + 2\zeta s + 1}$ with a Unit Step Input

## CHAPTER V

## FREQUENCY-WEIGHTED MEAN SQUARE ERROR

Ruchkin (Ref. 32) has modified the mean square error performance measure to deal with the case when error power is more objectionable in certain frequency bands than in others. This modification consists of the introduction of a weighting function into the expression for mean square error. Instead of

$$\text{minimum } \overline{\epsilon^2} \;=\; \text{minimum } \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_{\epsilon\epsilon}(\omega)\,d\omega \qquad (191)$$

Ruchkin proposes the criterion

$$\text{minimum EMS}_W \;=\; \text{minimum } \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_{\epsilon\epsilon}(\omega)W(\omega)\,d\omega \qquad (192)$$

where $W(\omega)$ is the frequency weighting function
$\Phi_{\epsilon\epsilon}(\omega)$ is the error power spectrum

Sections A and B of this chapter paraphrase Ruchkin's results, showing that the optimum free-form system is simply related to the Wiener system and establishing necessary restrictions upon the choice of $W(\omega)$. Section C extends Ruchkin's analysis to show that the criterion

$$\text{minimum } \left( \alpha_0 \overline{\epsilon^2} + \alpha_1 \overline{\left(\frac{d\epsilon}{dt}\right)^2} + \alpha_2 \overline{\left(\frac{d^2\epsilon}{dt^2}\right)^2} + \cdots \right) \qquad (193)$$

can also be expressed as a frequency-weighted mean square error. Section D considers the optimization of fixed-form systems. The effect of the choice of $W(\omega)$ upon the selectivity and validity of the $\text{EMS}_W$ criterion is also investigated in Sections B and C.

105

## A. DERIVATION OF THE MINIMUM EMS$_W$ SYSTEM

It will now be demonstrated that the minimum EMS$_W$ system can be obtained in an analogous fashion to the Wiener system derivation if $W(\omega)$, the frequency weighting in Eq 192, has certain properties in common with the signal noise and input power spectra. In the most general case, the input is the sum of correlated signal and noise. Hence,

$$\Phi_{rr} = \Phi_{ss} + \Phi_{nn} + \Phi_{sn} + \Phi_{ns} \tag{194}$$

$$\text{where} \quad \Phi_{ss} = \Phi_{ss}(\omega), \text{ signal spectrum}$$

$$\Phi_{nn} = \Phi_{nn}(\omega), \text{ noise spectrum}$$

$$\Phi_{sn} = \Phi_{sn}(j\omega) = \Phi_{ns}(-j\omega), \text{ cross-}$$
$$\text{spectrum between signal and noise}$$

It is assumed that $\Phi_{ss}$, $\Phi_{nn}$, and $\Phi_{rr}$ have the usual Hopf-Wiener factorization property, i.e., they are even rational functions of $\omega$, and therefore,

$$\Phi(\omega) = \Phi^+(j\omega)\Phi^-(j\omega) = |\Phi^+(j\omega)|^2 \tag{195}$$

$$\text{where} \quad \Phi^+(j\omega) \quad \text{has poles and zeros only in the left half-plane}$$

$$\Phi^-(j\omega) \quad \text{has poles and zeros only in the right half-plane}$$

$$\Phi^-(j\omega) = \Phi^+(-j\omega)$$

$W(\omega)$ is also assumed to be Hopf-Wiener factorable. Therefore, EMS$_W$ can be obtained by filtering the error signal with a system having the transfer function $W^+(j\omega)$ (as shown in Fig. 31) and evaluating $\overline{\epsilon_W^2}$.
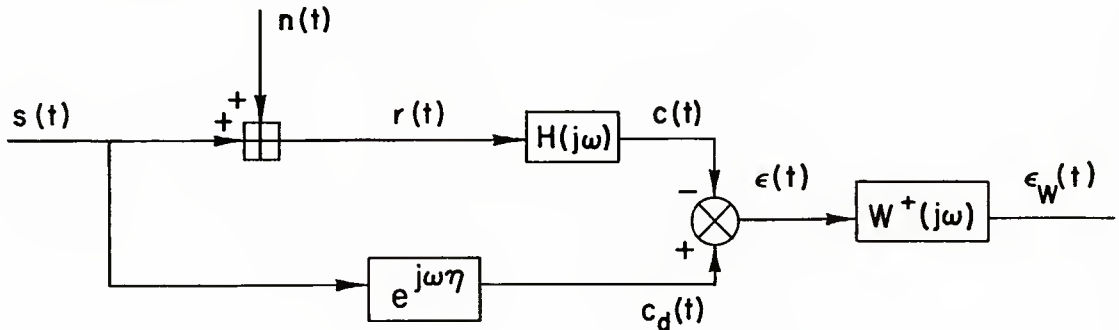


Figure 31. Block Diagram of System with Signal Plus Noise Input

From Fig. 31 it can be seen that the frequency-weighted mean square error is

$$\text{EMS}_W = \overline{\epsilon_W^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_{\epsilon_W\epsilon_W}(\omega)d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_{\epsilon\epsilon}(\omega)W(\omega)d\omega \qquad (196)$$

where $\quad W(\omega) = W^+(j\omega)W^+(-j\omega)$

and $\quad \Phi_{\epsilon_W\epsilon_W}(\omega) =$ the spectrum of $\epsilon_W$, the weighted error

The error power spectrum expressed in terms of the signal, noise, and cross-spectra is the same as in the Wiener problem. From Ref. 50, p. 239,

$$\Phi_{\epsilon\epsilon}(\omega) = \left|e^{j\omega\eta} - H\right|^2\Phi_{ss} + |H|^2\Phi_{nn}$$

$$- \left[e^{-j\omega\eta} - H^*\right]H\Phi_{sn} - \left[e^{j\omega\eta} - H\right]H^*\Phi_{ns} \qquad (197)$$

where $\quad \eta$ is the prediction time

$\quad H = H(j\omega)$ is a general system transfer function

$\quad H^*(j\omega) = H(-j\omega)$

The system that minimizes the frequency-weighted mean square error can be found by using the formula for the Wiener optimum linear system which is (as shown in Chapter I)

$$H_W(j\omega) = \frac{1}{\Phi_{rr}^+}\left[e^{j\omega\eta}\left(\frac{\Phi_{ss} + \Phi_{ns}}{\Phi_{rr}^-}\right)\right]_+ \qquad (198)$$

If, in Eq 197, $\Phi_{ss}$, $\Phi_{nn}$, $\Phi_{sn}$, and $\Phi_{ns}$ are replaced with $W\Phi_{ss}$, $W\Phi_{nn}$, $W\Phi_{sn}$, and $W\Phi_{ns}$, respectively, then the left-hand side becomes $W\Phi_{\epsilon\epsilon}$ ($\equiv\Phi_{\epsilon_W\epsilon_W}$). With these substitutions, the problem of minimizing EMS, Eq 196, can be regarded as that of minimizing the mean square error <u>using</u> <u>the</u> <u>modified</u> <u>input</u> <u>spectrum</u> $\underline{W\Phi_{rr}}$ <u>given</u> <u>above</u>. Inserting these modified spectra into Eq 198 yields the optimum $\text{EMS}_W$ linear system, $H_{\text{EMS}_W}(j\omega)$, as

$$H_{\text{EMS}_W}(j\omega) = \frac{1}{\Phi_{rr}^+W^+}\left[\frac{e^{j\omega\eta}(\Phi_{ss} + \Phi_{ns})W^+W^-}{\Phi_{rr}^-W^-}\right]_+ \qquad (199)$$

which can be simplified to the result given by Ruchkin (Ref. 32):

$$H_{EMS_W}(j\omega) \quad = \quad \frac{1}{\Phi_{rr}^{+}W^{+}} \left[ \frac{e^{j\omega\eta}(\Phi_{ss} + \Phi_{ns})W^{+}}{\Phi_{rr}^{-}} \right]_{+} \tag{200}$$

For $W(\omega) \equiv 1$ the optimum $EMS_W$ system becomes the Wiener system. Ruchkin does not consider the characteristics of the optimum $EMS_W$ system and proceeds to discuss the necessary restrictions upon $W(\omega)$ that are implicit in Eq 200. However, it requires only a brief digression to show that the optimum $EMS_W$ system is simply related to the Wiener system. This will now be proved before continuing to discuss the limitations on $W(\omega)$. It will be shown that the optimum $EMS_W$ system can be represented by two systems in parallel, one of which is the Wiener system for the given signal and noise. For the time being it will be assumed that $W(\omega)$ has an equal number of zeros and poles. The reasons for this assumption are made clear in Section B.

Equation 200 can be written as

$$H_{EMS_W} \quad = \quad \frac{\left[(A + B)W^{+}\right]_{+}}{\Phi_{rr}^{+}W^{+}} \tag{201}$$

where $A = A(j\omega)$ is that part of the partial fraction expansion of $e^{j\omega\eta}\left(\dfrac{\Phi_{ss} + \Phi_{ns}}{\Phi_{rr}^{-}}\right)$ with only left-half-plane poles

$B = B(j\omega)$ is that part of the partial fraction expansion of $e^{j\omega\eta}\left(\dfrac{\Phi_{ss} + \Phi_{ns}}{\Phi_{rr}^{-}}\right)$ with only right-half-plane poles

Equation 201 can be rewritten as

$$H_{EMS_W}(j\omega) \quad = \quad \frac{\left[AW^{+}\right]_{+} + \left[BW^{+}\right]_{+}}{\Phi_{rr}^{+}W^{+}} \tag{202}$$

Since $AW^{+}$ has only left-half-plane terms, $\left[AW^{+}\right]_{+} = AW^{+}$ and Eq 202 becomes

108

$$H_{EMS_W}(j\omega) \;=\; \frac{A}{\Phi_{rr}^{+}} \;+\; \frac{[BW^{+}]_{+}}{\Phi_{rr}^{+}W^{+}} \tag{203}$$

and this can be written in terms of the Wiener system as

$$H_{EMS_W} \;=\; H_W + H_X \tag{204}$$

where $\quad H_W \;=\; \dfrac{A}{\Phi_{rr}^{+}} \qquad$ the Wiener system

$$H_X \;=\; \frac{[BW^{+}]_{+}}{\Phi_{rr}^{+}W^{+}} \qquad \begin{array}{l}\text{the system in parallel}\\ \text{with } H_W\end{array}$$

A and B are as in Eq 201

Thus, $H_{EMS_W}$ can be expressed as two systems in parallel, one of which is simply the Wiener system.

## Example of Optimum Frequency-Weighted Mean Square Error System

As an example to illustrate the foregoing, consider the following uncorrelated signal and noise input spectra:

$$\Phi_{ss} \;=\; \frac{1}{1+\omega^2} \;=\; \left.\frac{1}{1-s^2}\right|_{s=j\omega} \;=\; \frac{1}{(1+s)(1-s)} \tag{205}$$

$$\Phi_{nn} \;=\; N^2$$

$$\Phi_{rr} \;=\; \Phi_{ss} + \Phi_{nn} \;=\; \frac{1}{1-s^2} + N^2 \;=\; \frac{1+N^2-N^2s^2}{1-s^2}$$

And since $\Phi_{rr}$ is Hopf-Wiener factorable,

$$\Phi_{rr}^{+} \;=\; \frac{\sqrt{N^2+1}+Ns}{1+s}$$

$$\Phi_{rr}^{-} \;=\; \frac{\sqrt{N^2+1}-Ns}{1-s} \tag{206}$$

$H_{EMS_W}$ is given by Eq 200. For $\eta = 0$,

$$H_{EMS_W} = \frac{\left[\dfrac{\Phi_{ss}W^+}{\Phi_{rr}^-}\right]_+}{\Phi_{rr}^+W^+} = \frac{\left[\dfrac{W^+}{N(1+s)(A-s)}\right]_+}{\dfrac{A+s}{1+s}W^+} \tag{207}$$

$$\text{where} \quad A = \frac{\sqrt{N^2+1}}{N}$$

Expanding $\Phi_{ss}/\Phi_{rr}^-$ in partial fractions yields

$$\frac{\Phi_{ss}}{\Phi_{rr}^-} = \frac{1}{N(A+1)}\left(\frac{1}{1+s} + \frac{1}{A-s}\right) \tag{208}$$

Therefore, from Eq 201, $A(s)$ and $B(s)$ are

$$A(s) = \frac{1}{N(A+1)(1+s)}$$

$$B(s) = \frac{1}{N(A+1)(A-s)} \tag{209}$$

Substituting Eq 206 and 209 into Eq 203 for $H_{EMS_W}(s)$ yields

$$H_{EMS_W}(s) = \frac{1}{N^2(A+1)(A+s)} + \frac{\left[\dfrac{W^+}{N(A+1)(A-s)}\right]_+}{N\left(\dfrac{A+s}{1+s}\right)W^+} \tag{210}$$

where the first term is the Wiener system for the given signal and noise.

B. RESTRICTIONS UPON THE CHOICE OF THE FREQUENCY WEIGHTING FUNCTION

The restrictions upon the choice of the frequency weighting function will now be illustrated by continuing the above example. Assume the following form for $W^+(s)$:

$$W^+(s) \;=\; \frac{1 + \alpha s}{1 + \beta s} \tag{211}$$

Inserting this into the second term of Eq 210 yields (using Eq 204)

$$H_X(s) \;=\; \frac{\left[\dfrac{1 + \alpha s}{N(A + 1)(A - s)(1 + \beta s)}\right]_+}{\dfrac{N(A + s)}{1 + s}\left(\dfrac{1 + \alpha s}{1 + \beta s}\right)} \;=\; \frac{\dfrac{(1 - \frac{\alpha}{\beta})}{N(A + 1)(A + \frac{1}{\beta})(1 + \beta s)}}{\dfrac{N(A + s)}{1 + s}\left(\dfrac{1 + \alpha s}{1 + \beta s}\right)} \tag{212}$$

which can be simplified to

$$H_X(s) \;=\; \frac{(1 - \frac{\alpha}{\beta})(1 + s)}{N^2(A + 1)(A + \frac{1}{\beta})(A + s)(1 + \alpha s)} \tag{213}$$

Ruchkin states that $W^+(s)$ must have at least as many zeros as poles. A simple demonstration of this is obtained by letting $\alpha \rightarrow 0$ in Eq 211 and inspecting the final result. From Eq 213, $H_X(s)$ then becomes

$$H_X(s) \;=\; \frac{1 + s}{N^2(A + 1)(A + \frac{1}{\beta})(A + s)} \tag{214}$$

This system has a finite gain at all frequencies and, since the noise part of the input has infinite bandwidth, the output power will be infinite; clearly an undesirable result. It can be shown that if $W^+(s)$ had been a ratio of second-order polynomials in s and the above procedure followed, then $H_X(s)$ would have one more zero than pole. Therefore, it would be physically unrealizable in addition to having infinite output power. Hence, the weighting function must have at least as many zeros as poles. To see that it may have more zeros than poles, allow $\beta \rightarrow 0$ in $H_X(s)$, Eq 213 which yields

111

$$H_X(s) \;\; = \;\; \frac{-\alpha(1 + s)}{N^2(A + 1)(A + s)(1 + \alpha s)} \tag{215}$$

and therefore $H_{EMS_W}(s)$, Eq 204, becomes

$$H_{EMS_W}(s) \;\; = \;\; \frac{1}{N^2(A + 1)(A + s)} - \frac{\alpha(1 + s)}{N^2(A + 1)(A + s)(1 + \alpha s)} \tag{216}$$

which could be combined to yield

$$H_{EMS_W}(s) \;\; = \;\; \frac{1 - \alpha}{N^2(A + 1)(A + s)(1 + \alpha s)} \tag{217}$$

C.  FREQUENCY WEIGHTING AS A MEANS OF REPRESENTING A GENERAL
QUADRATIC CRITERION INVOLVING DERIVATIVES OF ERROR

The criterion

$$\text{minimum } (\alpha_0 \overline{\epsilon^2} + \alpha_2 \overline{\dot{\epsilon}^2} + \alpha_2 \overline{\ddot{\epsilon}^2} + \cdots) \tag{218}$$

has been suggested (in a more general quadratic form) by Bellman (Ref. 55) and by
Kalman and Koepcke (Ref. 56) for dynamic programming optimization techniques.
This is easily interpreted as a frequency-weighted form

$$EMS_W \;\; = \;\; \frac{1}{2\pi} \int_{-\infty}^{\infty} \underbrace{(\alpha_1^2 + \alpha_2^2 \omega^2 + \alpha_3^2 \omega^4 + \cdots)}_{W(\omega)} \Phi_{\epsilon\epsilon}(\omega) d\omega \tag{219}$$

where use has been made of the following relationships:

$$\overline{\dot{\epsilon}^2} \;\; = \;\; \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_{\dot{\epsilon}\dot{\epsilon}}(\omega) d\omega \;\; = \;\; \frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^2 \Phi_{\epsilon\epsilon}(\omega) d\omega$$

$$\tag{220}$$

$$\overline{\ddot{\epsilon}^2} \;\; = \;\; \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_{\ddot{\epsilon}\ddot{\epsilon}}(\omega) d\omega \;\; = \;\; \frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^4 \Phi_{\epsilon\epsilon}(\omega) d\omega$$

$$\vdots$$

The formula for the optimum linear system, given by Eq 204, still applies. An example of optimization with this criterion is given in the previous section (see Eq 215, 216, and 217) where it can be seen that the weighting function Eq 211 for $\beta \to 0$ is

$$W^+(s) = \lim_{\beta \to 0} \frac{1 + \alpha s}{1 + \beta s} = 1 + \alpha s \tag{221}$$

therefore
$$W(\omega) = 1 + \alpha^2 \omega^2 \tag{222}$$

and therefore in this case

$$EMS_W = \overline{\epsilon^2} + \alpha^2 \overline{\dot{\epsilon}^2} \tag{223}$$

The validity, selectivity, and ease of application of this criterion will be tested in Section D, where it will be applied to a second-order unit numerator system $H(s) = \dfrac{1}{s^2 + 2\zeta s + 1}$, where $\zeta$ is the variable parameter.

## D. FIXED-FORM (PHILLIPS-TYPE) OPTIMIZATION OF $EMS_W$

This section derives the optimum fixed-form (Phillips) system of the form $H(s) = \dfrac{1}{s^2 + 2\zeta s + 1}$ for two different frequency weighting functions:

$$W_a(\omega) = \lim_{\beta \to 0} \left| \frac{1 + \alpha s}{1 + \beta s} \right|^2 \to (1 + \alpha^2 \omega^2) \tag{224}$$

$$W_b(\omega) = \left| \frac{s^2 + 2\zeta_1 s + 1}{s^2 + 2\zeta_2 s + 1} \right|^2 \tag{225}$$

The reason for the limiting process in $W_a(\omega)$ will become clear in the discussion. It will be shown that neither $W_a$ nor $W_b$ yields a valid and selective criterion.

As demonstrated in the previous section, inserting the weighting function $W_a(\omega)$ into $EMS_W$ yields the simplest form of Bellman's performance measure:

$$EMS_W = \overline{\epsilon^2} + \alpha^2 \overline{\dot{\epsilon}^2} \tag{226}$$

113

$EMS_W$ (Eq 192) could be evaluated on an analog computer using the circuit in Fig. 31 where $W^+(j\omega)$ is

$$W^+(j\omega) = 1 + \alpha s \qquad (227)$$

But even with operational amplifiers this is not strictly realizable, since any differentiating circuit will fall off at very high frequencies. Therefore, the use of

$$W_a^+(j\omega) = \frac{1 + \alpha s}{1 + \beta s} \qquad (228)$$

where $\beta$ is much smaller than any of the time constants of the error signal would be a more realistic and practical system. It will be shown that the optimization of $H(s) = \dfrac{1}{s^2 + 2\zeta s + 1}$ for an integrated white spectrum desired input signal (zero noise) is essentially independent of $\beta$ if $\beta$ is sufficiently small.

The terms comprising $\Phi_{\epsilon\epsilon}(\omega)$ (Eq 197) are therefore

$$\Phi_{ss} = \frac{1}{\omega^2} \qquad (229)$$

$$\Phi_{nn} = \Phi_{ns} = 0$$

$$H(s) = \frac{1}{s^2 + 2\zeta s + 1}$$

$$\eta = 0$$

and therefore the integrand of $EMS_W$ becomes

$$\Phi_{\epsilon\epsilon}W = \left| \left( \frac{s + 2\zeta}{s^2 + 2\zeta s + 1} \right) \left( \frac{1 + \alpha s}{1 + \beta s} \right) \right|^2_{s=j\omega} \qquad (230)$$

Evaluating $EMS_W$ using the tables in Ref. 6 yields

$$EMS_W = \left[ \frac{\alpha^2(2\zeta + \beta) + [(2\zeta\alpha + 1)^2 - 4\zeta\alpha]\beta + 4\zeta^2(2\zeta\beta + 1)\beta}{2\beta[(2\zeta + \beta)(2\beta\zeta + 1) - \beta]} \right] \qquad (231)$$

Dividing numerator and denominator by $\beta$ and simplifying, this becomes

$$EMS_W = \left[\frac{(4\zeta^2 + 1)(\alpha^2 + 1) + \frac{2\alpha^2\zeta}{\beta} + 8\zeta^3\beta}{4\zeta(1 + 2\zeta\beta + \beta^2)}\right] \qquad (232)$$

For small $\beta$ the last term in the numerator is negligible and Eq 232 becomes approximately

$$EMS_W \doteq (\zeta + \frac{1}{4\zeta})(\alpha^2 + 1) + \left(\frac{\alpha^2}{2\beta(1 + 2\zeta\beta + \beta^2)}\right) \qquad (233)$$

Expanding the last term in a Taylor series about $\beta = 0$ yields

$$EMS_W \doteq (\zeta + \frac{1}{4\zeta})(\alpha^2 + 1) + \frac{\alpha^2}{2\beta}\left\{1 - 2\zeta\beta + \beta^2(4\zeta^2 - 1) \cdots\right\} \qquad (234)$$

Expanding Eq 234 yields

$$EMS_W \doteq (\zeta + \frac{1}{4\zeta})(\alpha^2 + 1) + \frac{\alpha^2}{2\beta} - \zeta\alpha^2 + \alpha^2(4\zeta^2 - 1)\beta - \cdots \qquad (235)$$

and for small $\beta$ the frequency-weighted mean square error finally becomes

$$EMS_W \doteq \zeta + \frac{1 + \alpha^2}{4\zeta} + \frac{\alpha^2}{2\beta} \qquad (236)$$

The value of $\zeta$ that minimizes $EMS_W$ is found by differentiating Eq 236 with respect to $\zeta$.

$$\frac{dEMS_W}{d\zeta} \doteq 1 - \frac{1 + \alpha^2}{4\zeta^2} = 0 \qquad (237)$$

and solving for $\zeta$ yields

$$\zeta = \frac{\sqrt{1 + \alpha^2}}{2} \qquad (238)$$

For this criterion to be valid ($\zeta = 0.7$) requires an $\alpha$ of about one. Eq 236 is graphed in Fig. 32 for $\alpha = 1$, $\beta = 1$. Fig. 32 shows that the selectivity of this $EMS_W$ criterion is no better than that of $\overline{\epsilon^2} = \zeta + \frac{1}{4\zeta}$.
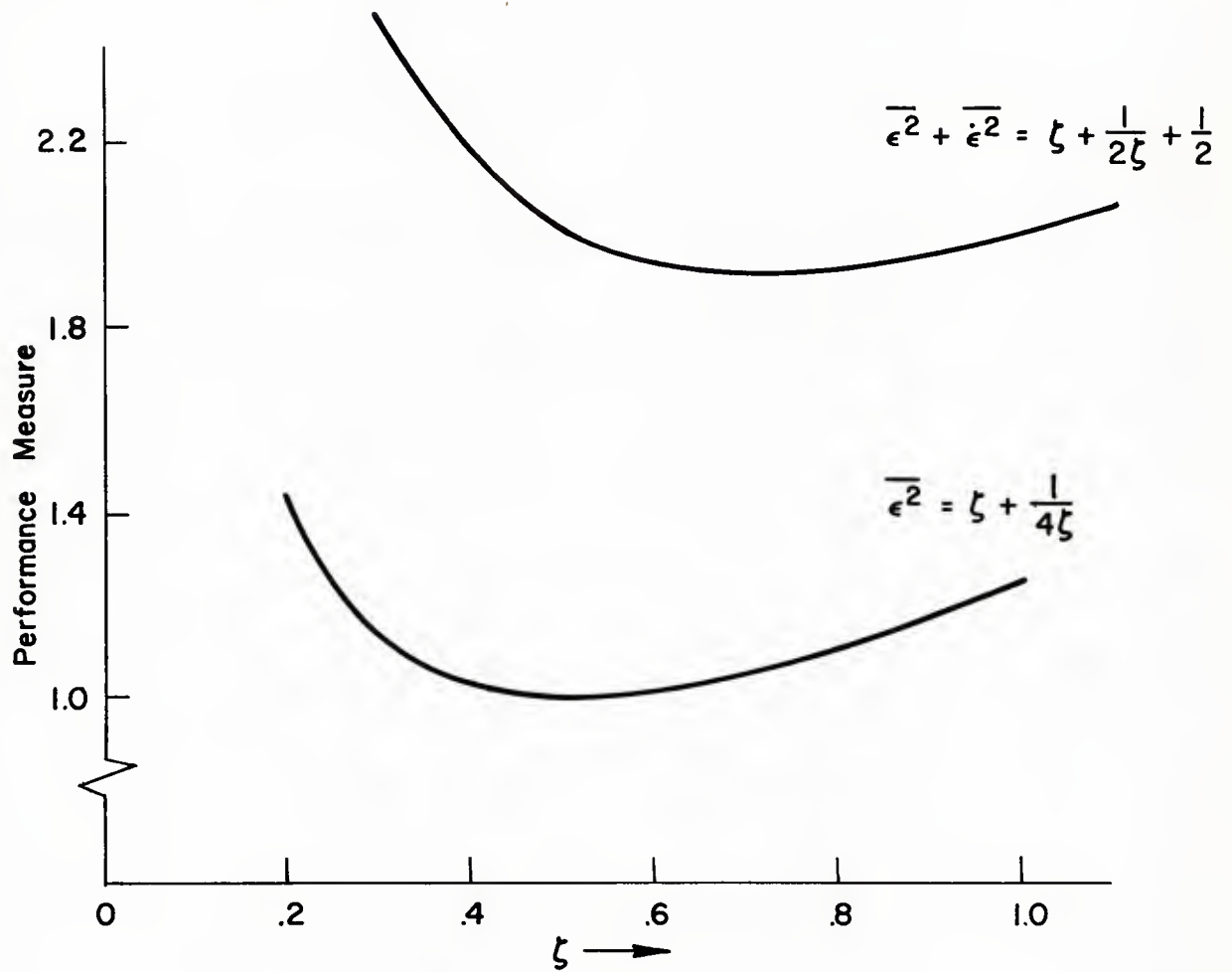
115

Figure 32. Comparison of $\overline{\epsilon^2}$ and $\overline{\epsilon^2} + \overline{\dot{\epsilon}^2}$ for $H(s) = \dfrac{1}{s^2 + 2\zeta s + 1}$

with an Integrated White Spectrum Input $\Phi_{ss}(\omega) = \dfrac{1}{\omega^2}$

116

As an example of optimization with the second weighting function, Eq 225,

$$W_b(\omega) = \left| \frac{s^2 + 2\zeta_1 s + 1}{s^2 + 2\zeta_2 s + 1} \right|^2_{s=j\omega} \tag{239}$$

the same system, $H(s)$, and input of the previous example will be used. For $\zeta_1 \doteq 1$ and $\zeta_2 \ll 1$, this weighting function will tend to accentuate the mid-frequency components of $\Phi_{\epsilon\epsilon}$ which depend mainly on $\zeta$. One would therefore expect the criterion to be highly selective for these values of $\zeta_1$ and $\zeta_2$. Therefore,

$$\Phi_{\epsilon\epsilon}W = \left| \left(\frac{s + 2\zeta}{s^2 + 2\zeta s + 1}\right)\left(\frac{s^2 + 2\zeta_1 s + 1}{s^2 + 2\zeta_2 s + 1}\right) \right|^2_{s=j\omega} \tag{240}$$

Using the table of Ref. 6, $EMS_W$ becomes

$$EMS_W = \frac{(4\zeta^2 + 1)(\zeta\zeta_2 + \zeta_1^2)}{4\zeta(\zeta + \zeta_2)\zeta_2} \tag{241}$$

This is plotted in Fig. 33 as a function of $\zeta$ for various values of $\zeta_1$ and $\zeta_2$. In addition, plots of ITAE, $ITE^2$, and $IT^2E^2$ for a step input to the same system are reproduced from Ref. 1 for selectivity comparisons. The analytic forms of these performance measures are rational functions of $\zeta$ and therefore they have been plotted "Bode fashion," taking advantage of their asymptotic character (except for ITAE which contains transcendental terms). It must be remembered though that only positive real values of $\zeta$ are considered in the "Bode plots."

The $\overline{\epsilon^2}$ can be obtained from $EMS_W$ when $\zeta_1 = \zeta_2$ and is graphed in Fig. 33 as curve ①. The selectivity is poor. The steepest possible asymptotes obtainable using the weighting function of Eq 239 and the actual $EMS_W$ are shown as curve ②. The selectivity is little better than that of the mean square error. In comparison, "good" selectivity is indicated by curves ③, ④, and ⑤ for ITAE, $IT^2E^2$, and $ITE^2$, respectively.
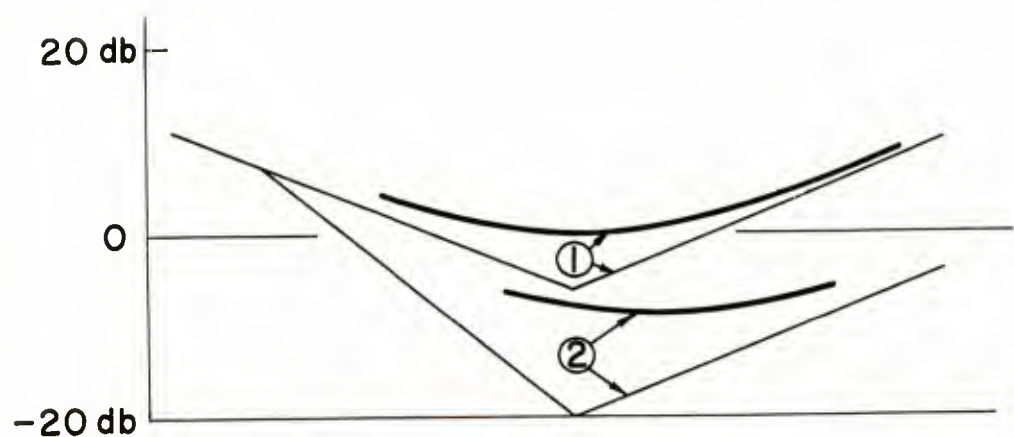
$$EMS_W = \frac{(4\zeta^2 + 1)(\zeta\zeta_2 + \zeta_1^2)}{4\zeta\zeta_2(\zeta + \zeta_2)}$$

① $\zeta_2 = \zeta_1$

② $\zeta_2 = 0.1$

$\zeta_1 = 0.224$
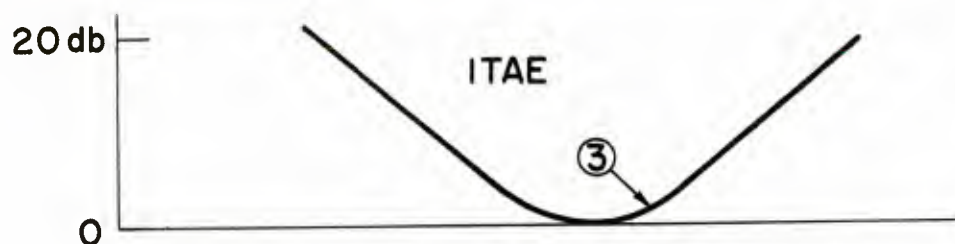
③ ITAE (formula given in Ref 1)
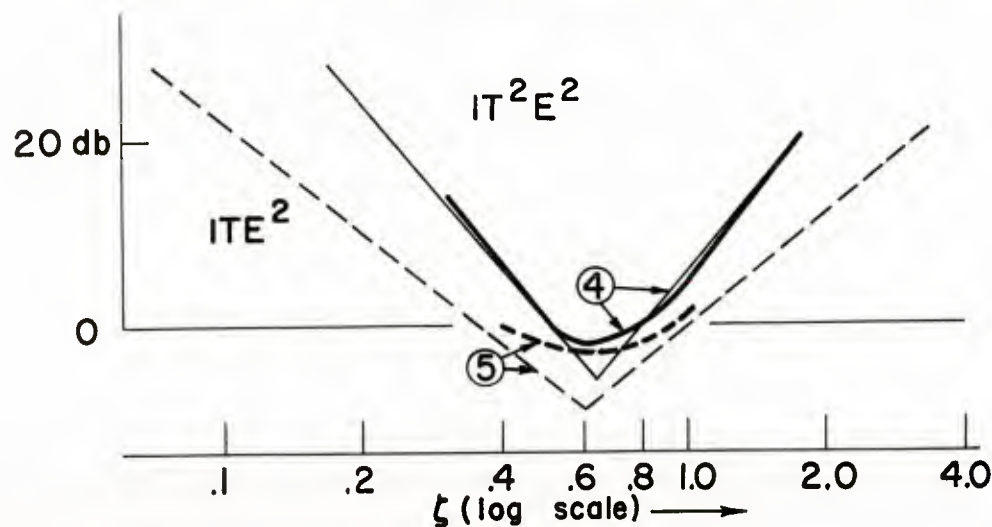
—— ④ $IT^2E^2 = 2\zeta^3 - \frac{\zeta}{2} + \frac{1}{8\zeta} + \frac{1}{8\zeta^3}$

--- ⑤ $ITE^2 = \zeta^2 + \frac{1}{8\zeta^2}$

Figure 33. Selectivity Comparisons of $EMS_W$, $\overline{\epsilon^2}$, ITAE, $ITE^2$, and $IT^2E^2$

CHAPTER VI

A NOTE ON COMBINED AND CONSTRAINED CRITERIA, AND CONCLUSIONS

The term "combined criteria" is used to denote performance criteria that are functions of more than one dependent variable, e.g.,

The generalized
error function,     minimum G.E.F. $\equiv f(\epsilon, \tau)$

Glover's criterion, minimum P.E. = minimum $f(\epsilon, c, \delta)$

Combined criteria can sometimes be expressed as frequency-weighted forms, e.g., minimum $\overline{\epsilon^2} + \kappa\overline{\dot{\epsilon}^2}$ is equivalent to minimum $\frac{1}{2\pi} \int_{-\infty}^{\infty} W(\omega)\Phi_{\epsilon\epsilon}(\omega)d\omega$, where $W(\omega)$ is a frequency-weighting factor incorporating a constant $\kappa$ which expresses the relative importance of error and time rate of change of error quantities as measures of the dynamic performance of a given system.

In order to put combined criteria into concrete usable forms, the linking constants or equivalent frequency-weighting functions must be chosen to reflect system requirements. The present generalized study can offer little advice on the choice of these weighting parameters; in Chapters III, IV, and V we have attempted to illustrate the consequences that follow once certain choices of these parameters are made. The reader must decide for himself whether these choices are appropriate for the particular system that he is studying.

Combined criteria are frequently expressed in the form "minimum $\overline{f(\epsilon)}$" with a constraint on some other parameter, such as bandwidth or peak power. Such constrained criteria have been extensively studied by Westcott (Ref. 11), Newton, Gould, and Kaiser (Ref. 6), Chang (Ref. 10), Hung (Ref. 9), and others. Westcott noted that constrained criteria could be divided into two classes:

Class 1:   A condition is imposed upon the system transfer
           function either at a particular frequency or as
           a trend with frequency.

Class 2:   The constraint is expressible as an integral
           over the range of the function being varied and
           is required to have a constant value or to be
           less than some specified value, e.g., limiting
           of total energy.

Reference 11 and most subsequent analyses of random input performance criteria

deal with Class 2 constraints. The difficulty in selecting reasonable values of
the weighting parameters to assess combined criteria becomes aggravated when
constraints are introduced. Mathematical procedures required to obtain the
optimum system have been detailed in Ref. 6 and 10. In the present report only
a brief note will be given to indicate the techniques that are available; details
of the computational procedures can be found in the references cited.

Newton, Gould, and Kaiser consider constrained criteria of the form
"minimum $\overline{\epsilon^2}$," with a constraint upon the rms input to a supposedly nonlinear
component of the system. The object is to optimize the system subject to the
restriction that the nonlinear element must operate in the linear (unsaturated)
part of its range. The analytical procedure adopted consists of the replacement
of the nonlinear element by a substitute linear component having identical char-
acteristics to those of the actual component over the linear part of its operating
range. The rms input to the substituted linear element is then constrained to be
less than that maximum input amplitude for which the actual system remains linear.
It is shown in Ref. 6 that this procedure is equivalent to minimizing the proba-
bility of saturation, provided that that probability is small. Lagrange
multipliers are used to incorporate the constraint into the minimization procedure.

Related techniques are used in Ref. 6, p. 215, to minimize mean square error
subject to a constraint on bandwidth, although the term "bandwidth" is used only
in a rather general sense as "the frequency over which the system will have an
output approximately equal to the desired output." Such a definition is too loose
to be incorporated directly into standard variational calculus or other minimiza-
tion procedures, and the bandwidth is therefore calibrated in terms of the mean
square noise transmitted by a specified standard filter. The mean square error
is then minimized with this "calibrated constraint" being taken in account by
means of Lagrange multipliers. Ref. 6 presents examples of this optimization
procedure for deterministic inputs only; however, as noted in Ref. 9, the modifi-
cations required to deal with stationary random inputs are easily made. In
Ref. 9 the bandwidth of the input to a specified component is constrained to
avoid exciting undesirable high-frequency modes associated with that component.
In flight control system design this technique could be useful in minimizing mean
square error without exciting aeroelastic or autopilot modes.

The computational labor involved in the analytic minimization of constrained performance measures is often formidable. This point is noted on p. 246 of Ref. 6, where it is suggested that the analytic minimization procedure should be used mainly as a guide for trial-and-error optimization procedures, "pilot" analytic minimization using simplified systems and inputs being made to approximately determine the lowest attainable value of the particular performance measure being considered. Some progress toward shortening this process has been made by Chang. His "root square locus" method (Ref. 10) goes some way toward combining the precision of analytical optimization techniques with the ease of iteration afforded by graphical procedures.

# CONCLUSIONS

1.  Performance criteria for linear constant-coefficient systems with random inputs have been investigated, with particular reference to flight control systems. The application of performance measures has been facilitated by substituting for the actual flight control system an "equivalent" low-order linearized system having similar dynamic characteristics. This equivalent system was constructed by dividing the actual system transfer function into regions of interest defined by

(a) $|G(j\omega)| \gg 1$, over which $\left|\dfrac{G(j\omega)}{1 + G(j\omega)}\right| \doteq 1$

(b) $|G(j\omega)| \ll 1$, over which $\left|\dfrac{G(j\omega)}{1 + G(j\omega)}\right| \doteq |G(j\omega)|$

(c) $|G(j\omega)| \doteq 1$

The form of $\left|\dfrac{G(j\omega)}{1 + G(j\omega)}\right|$ in the last region defines the dominant modes of the closed-loop system response, and can usually be closely approximated by a system of first, second, or third order.

2.  For stationary random inputs, many criteria are equivalent to minimum $\overline{\epsilon^2}$. This criterion is relatively easy to apply using "pencil and paper" techniques if use is made of the equivalent system concept. It possesses a compatible deterministic criterion (minimum $\int_{0}^{\infty} \epsilon^2 dt$); hence, operations upon random quantities required to compute $\overline{\epsilon^2}$ can be replaced by operations upon their more easily visualized transient analogs. Its principal disadvantages are that it yields rather lightly damped systems and is unselective, i.e., the $\overline{\epsilon^2}$ of the optimum system is only slightly less than that of a wide range of off-optimum systems.

3.  No criterion has been found that has the advantages of minimum $\overline{\epsilon^2}$ without comparable disadvantages.

4.  Of those criteria that did not reduce to minimum $\overline{\epsilon^2}$, the following deserve further investigation:

a. Exceedance criteria (for stationary Gaussian inputs):
Exceedances are relatively easy to calculate, but, in
most of the examples studied, gave very heavily damped
systems. The implications on dynamic performance of
the growing use of exceedances as fatigue criteria
requires consideration.

b. Frequency-weighted criteria: These may be suitable
where the frequency-weighting can be specified so that
it directly reflects performance requirements. Further
study is required to determine whether this is feasible
for a class of flight control systems sufficiently
broad to be of general interest.

5. Zaborszky and Diesel's probabilistic square error criterion is suitable
for nonstationary situations where the weighting function describing the varia-
tion of error importance with time can be specified.

6. A wide variety of criteria have been surveyed and many new results and
techniques have been produced. These include:

a. Proof that for stationary inputs Murphy and Bold's
time-weighted square error criterion reduces to
minimum $\overline{\epsilon^2}$.

b. Simplified procedures for evaluating the probabilistic
square error measure.

c. Generalized definition of the concept of compatible
criteria and exploitation of this concept to obtain
transient analogs for both uncorrelated and corre-
lated signal and noise.

d. A technique for applying Glover's amplitude-weighted
criterion to fixed-form systems.

REFERENCES

1. Wolkovitch, J., R. E. Magdaleno, D. T. McRuer, F. D. Graham, and
   J. D. McDonnell, <u>Performance Criteria for Linear Constant-Coefficient
   Systems with Deterministic Inputs</u>, ASD Technical Report 61-501,
   December 1961.

2. Truxal, J. G., <u>Automatic Feedback Control System Synthesis</u>, McGraw-Hill
   Book Company, Inc., New York, 1955.

3. James, H. M., N. B. Nichols, and R. S. Phillips, <u>Theory of Servomechanisms</u>,
   McGraw-Hill Book Company, Inc., New York, 1947.

4. Wiener, N., <u>The Extrapolation, Interpolation, and Smoothing of Stationary
   Time Series</u>, John Wiley and Sons, Inc., New York, 1949.

5. <u>Methods of Analysis and Synthesis of Piloted Aircraft Flight Control Systems</u>,
   BuAer Report AE-61-4-I, Northrop Aircraft, Inc., March 1952.

6. Newton, G. C., Jr., L. A. Gould, and J. F. Kaiser, <u>Analytical Design of
   Linear Feedback Controls</u>, John Wiley and Sons, Inc., New York, 1957.

7. Solodovnikov, V. V., <u>Introduction to the Statistical Dynamics of Automatic
   Control Systems</u>, Dover Publications, Inc., New York, 1960.

8. Laning, J., and R. H. Battin, <u>Random Processes in Automatic Control</u>,
   McGraw-Hill Book Company, Inc., New York, 1956.

9. Hung, J. C., "Control Systems with Minimum Spectral Bandwidth of Plant Input,"
   <u>IRE Trans.</u>, Vol. AC-6, February 1961.

10. Chang, S. S. L., <u>Synthesis of Optimum Control Systems</u>, McGraw-Hill Book
    Company, Inc., New York, 1961.

11. Westcott, J. H., "The Introduction of Constraints Into Feedback System Design,"
    <u>IRE Trans.</u>, Vol. CT-1, September 1954.

12. Zadeh, L. A., and J. R. Ragazzini, "An Extension of Wiener's Theory of
    Prediction," <u>J. Applied Physics</u>, July 1950.

13. Sherman, S., "Non-Mean-Square Error Criteria," <u>IRE Trans.</u>, Vol. IT-4,
    September 1958.

14. Pugachev, V. S., "The Determination of an Optimal System by Some Arbitrary
    Criterion," <u>Automation and Remote Control</u>, Vol. 19, No. 6, June 1958.

15. Brown, J. L., Jr., "Asymmetric Non-Mean-Square Error Criteria," <u>IRE Trans.</u>,
    Vol. AC-7, January 1962.

16. Kaufman, William M., <u>The Use of Generalized Error Criteria in the Statistical
    Evaluation of System Performance</u>, Ph.D. Thesis, College of Engineering
    and Science, Carnegie Institute of Technology, 1955.

17. Rice, S. O., "Mathematical Analysis of Random Noise," from Nelson Wax, ed., Selected Papers on Noise and Stochastic Processes, Dover Publications, Inc., New York, 1954.

18. Ragazzini, J. R., and L. A. Zadeh, "Probability Criterion for the Design of Servomechanisms," J. Applied Physics, February 1949.

19. Ruina, J. P., and M. E. Van Valkenburg, "Stochastic Analysis of Automatic Tracking Systems," from J. F. Coales, ed., Automatic and Remote Control (Proc. First International Congress of the IFAC, Moscow, 1960), Vol. II, Butterworth, Inc., Washington, D.C., 1960.

20. Murphy, G. J., and N. T. Bold, "Optimization Based on a Square-Error Criterion with an Arbitrary Weighting Function," IRE Trans., Vol. AC-5, January 1960.

21. Glover, C. C., "Optimum Prediction with a Mean Weighted Square Error Criterion," IRE Trans., Vol. AC-6, February 1961.

22. Zaborszky, J., and J. W. Diesel, "Probabilistic Error as a Measure of Control System Performance," AIEE Trans., Pt. II (Applications and Industry), Vol. 78, July 1959.

23. Zaborszky, J., and J. W. Diesel, "A Statistically Averaged Error Criterion for Feedback System Synthesis," J. Aero/Space Sciences, February 1960.

24. Zaborszky, J., and J. W. Diesel, "Design of Sampled Data Control Systems for Minimum Probabilistic Error," AIEE Trans., Pt. II (Applications and Industry), Vol. 79, May 1960.

25. Zaborszky, J., and J. W. Diesel, "Design of Continuous Linear Control Systems for Minimum Probabilistic Error," AIEE Trans., Pt. II (Applications and Industry), Vol. 79, May 1960.

26. Zaborszky, J., and J. W. Diesel, "Design for Minimum Probabilistic Error of Continuous Linear Control Systems Subject to Constraints," AIEE Trans., Pt. II (Applications and Industry), Vol. 79, May 1960.

27. Zaborszky, J., and J. W. Diesel, "Investigation of the Feasibility of Designing Homing Aircraft Flight Control Systems for Minimum Probabilistic Error," AIEE Trans., Pt. II (Applications and Industry), Vol. 79, May 1960.

28. Schultz, W. C., and V. C. Rideout, "A General Criterion for Servo Performance," Proc. National Electronics Conference, Vol. 13, 1957.

29. Benedict, T. R., and V. C. Rideout, "Error Determination for Optimum Predicting Filters," Proc. National Electronics Conference, Vol. 13, 1957.

30. Spooner, M. G., and V. C. Rideout, "Correlation Studies of Linear and Nonlinear Systems," Proc. National Electronics Conference, Vol. 12, 1956.

31. Schultz, W. C., and V. C. Rideout, "Control System Performance Measures: Past, Present and Future," IRE Trans., Vol. AC-6, February 1961.

32. Ruchkin, D. S., "A Frequency Weighted Mean-Square Error Criterion," IRE Trans., Vol. IT-7, July 1961.

33. Ule, L. A., "Weighted Least-Squares Smoothing Filters," IRE Trans., Vol. CT-2, June 1955.

34. Graham, Dunstan, and R. C. Lathrop, "The Synthesis of 'Optimum' Transient Response: Criteria and Standard Forms," AIEE Trans., Pt. II (Applications and Industry), Vol. 72, November 1953.

35. D'Azzo, J. J., and C. H. Houpis, Control System Analysis and Synthesis, McGraw-Hill Book Company, Inc., New York, 1960.

36. Westcott, J. H., "The Minimum-Moment-of-Error-Squared Criterion: A New Performance Criterion for Servos," Proc. IEE, London, 1954.

37. Cramer, H., Mathematical Methods of Statistics, Princeton University Press, Princeton, Ninth Printing, 1961.

38. Benedict, T. R., and M. M. Sondhi, "On a Property of Wiener Filters," Proc. IRE, July 1957.

39. Dwight, H. B., Tables of Integrals and Other Mathematical Data, The Macmillan Company, Third Ed., 1957.

40. Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," Trans. ASME (J. of Basic Engineering), March 1960.

41. Streets, R. B., Jr., Optimum Control Systems for a Class of Stochastic Signals, Massachusetts Institute of Technology Report 7967-R-4, March 1960.

42. Lubbock, J. K., "The Optimization of a Class of Non-Linear Filters," Proc. IEE, London, February 1959.

43. Press, H., M. T. Meadows, and I. Hadlock, A Re-evaluation of Data on Atmospheric Turbulence and Airplane Gust Loads for Application in Spectral Calculations, NACA Report 1272, 1956.

44. Etkin, B., Dynamics of Flight, John Wiley and Sons, Inc., New York, 1959.

45. Thorsen, K. R., and B. R. Bohne, "Application of Power Spectral Methods in Airplane and Missile Design," J. Aero/Space Sciences, February 1960.

46. Chernoff, M., and H. L. Rothman, "Unsteady Frequency Response Functions for Use in Power Spectral Analysis," J. Aero/Space Sciences, February 1962.

47. Press, H., and R. Steiner, An Approach to the Problem of Estimating Severe and Repeated Gust Loads for Missile Operations, NACA TN 4332, September 1958.

48. Schultz, W. C., Choice of Measures and Criteria for System Synthesis and Analysis, Ph.D. Dissertation, University of Wisconsin, 1958.

49. Zaborszky, J., and J. W. Diesel, "Optimization Based Upon a Square-Error Criterion with an Arbitrary Weighting Function," IRE Trans., Vol. AC-6, September 1961.

50. Davenport, W. B., Jr., and W. L. Root, Random Signals and Noise, McGraw-Hill Book Company, Inc., New York, 1958.

51. Mishkin, E., and L. Braun, Adaptive Control Systems, McGraw-Hill Book Company, Inc., New York, 1961.

52. Stone, C. R., A Mathematical Method for Optimizing the Gains of a Linear System Minneapolis-Honeywell Regulator Company Aero Report 49174-21 TR1, 15 February 1957.

53. Aigrain, P. R., and E. M. Williams, "Design of Optimum Transient Response Amplifiers," Proc. IRE, August 1949.

54. Lee, Y. W., and J. B. Wiesner, "Correlation Functions and Communications Applications," Electronics, Vol. 23, June 1950.

55. Bellman, R., Adaptive Control Processes: A Guided Tour, Princeton University Press, Princeton, 1961.

56. Kalman, R. E., and R. W. Koepcke, "Optimal Synthesis of Linear Sampling Control Systems Using Generalized Performance Indexes," Trans. ASME (J. of Basic Engineering), November 1958.

57. Durand, T. S., and H. R. Jex, Handling Qualities in Single-Loop Roll Tracking Tasks: Theory and Simulator Experiments, ASD Technical Documentary Report 62-507, May 1962.